



This is a digital copy of a book that was preserved for generations on library shelves before it was carefully scanned by Google as part of a project to make the world's books discoverable online.

It has survived long enough for the copyright to expire and the book to enter the public domain. A public domain book is one that was never subject to copyright or whose legal copyright term has expired. Whether a book is in the public domain may vary country to country. Public domain books are our gateways to the past, representing a wealth of history, culture and knowledge that's often difficult to discover.

Marks, notations and other marginalia present in the original volume will appear in this file - a reminder of this book's long journey from the publisher to a library and finally to you.

Usage guidelines

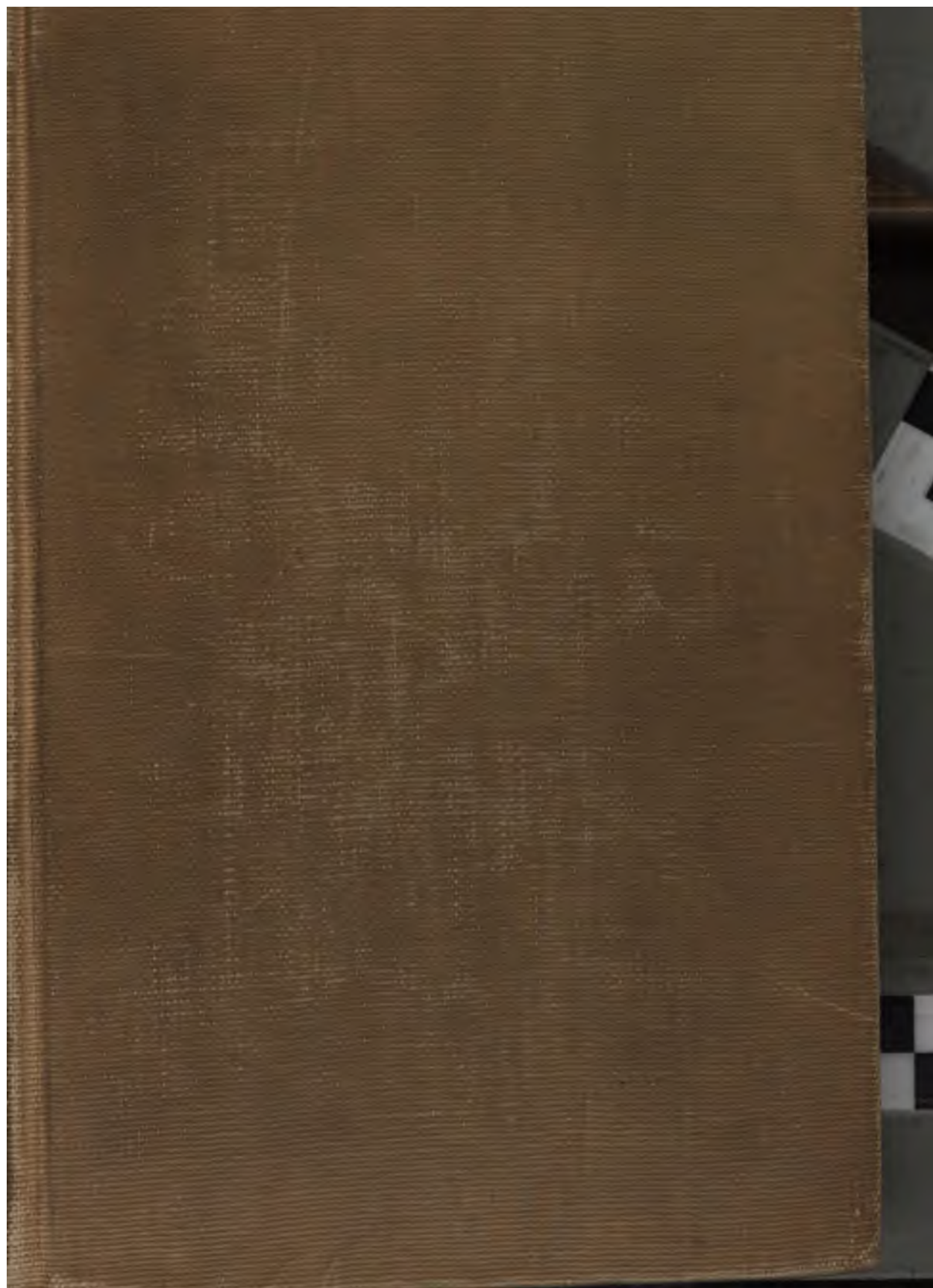
Google is proud to partner with libraries to digitize public domain materials and make them widely accessible. Public domain books belong to the public and we are merely their custodians. Nevertheless, this work is expensive, so in order to keep providing this resource, we have taken steps to prevent abuse by commercial parties, including placing technical restrictions on automated querying.

We also ask that you:

- + *Make non-commercial use of the files* We designed Google Book Search for use by individuals, and we request that you use these files for personal, non-commercial purposes.
- + *Refrain from automated querying* Do not send automated queries of any sort to Google's system: If you are conducting research on machine translation, optical character recognition or other areas where access to a large amount of text is helpful, please contact us. We encourage the use of public domain materials for these purposes and may be able to help.
- + *Maintain attribution* The Google "watermark" you see on each file is essential for informing people about this project and helping them find additional materials through Google Book Search. Please do not remove it.
- + *Keep it legal* Whatever your use, remember that you are responsible for ensuring that what you are doing is legal. Do not assume that just because we believe a book is in the public domain for users in the United States, that the work is also in the public domain for users in other countries. Whether a book is still in copyright varies from country to country, and we can't offer guidance on whether any specific use of any specific book is allowed. Please do not assume that a book's appearance in Google Book Search means it can be used in any manner anywhere in the world. Copyright infringement liability can be quite severe.

About Google Book Search

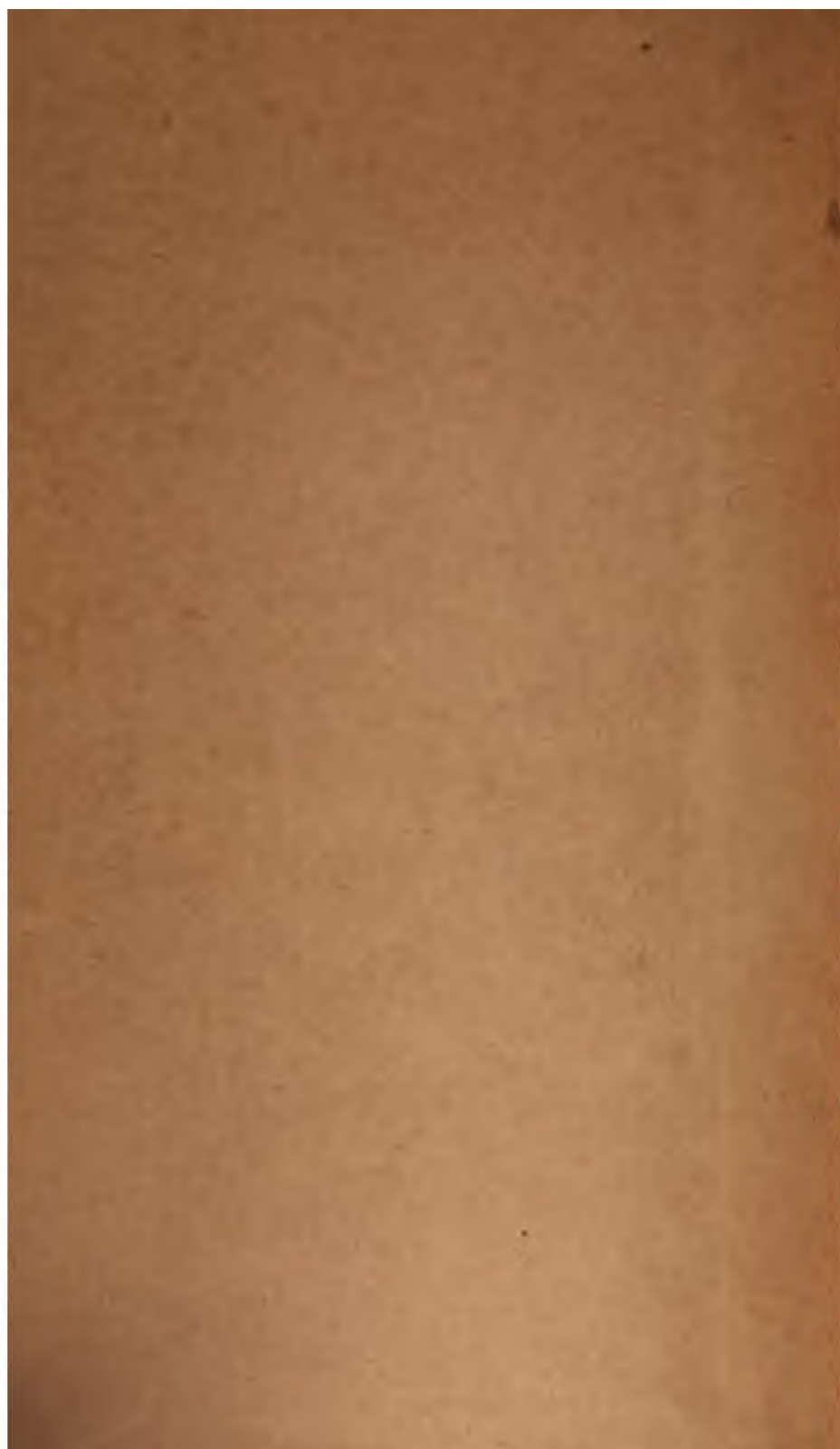
Google's mission is to organize the world's information and to make it universally accessible and useful. Google Book Search helps readers discover the world's books while helping authors and publishers reach new audiences. You can search through the full text of this book on the web at <http://books.google.com/>

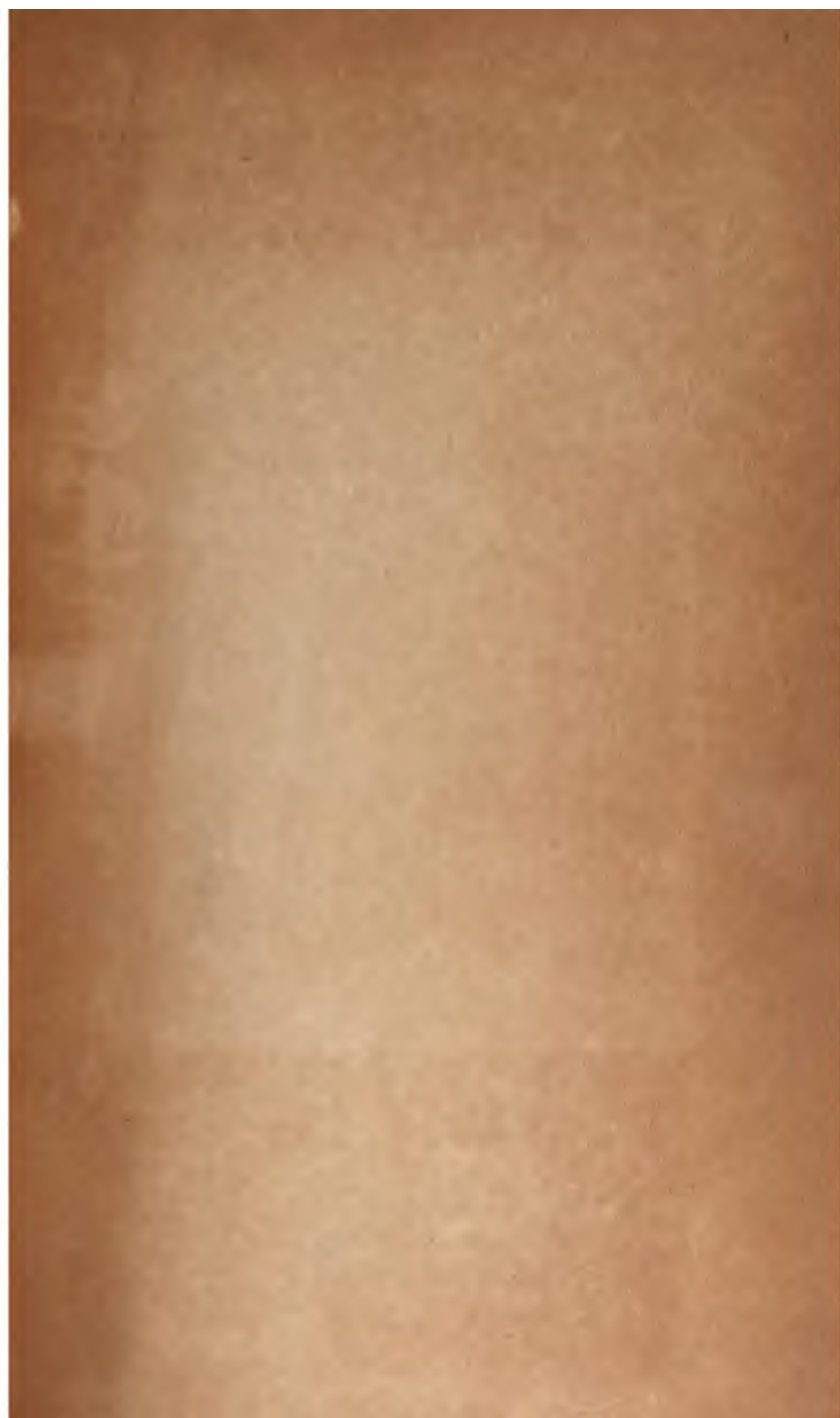




STANFORD·VNIVERSITY·LIBRARY







3 8 2-

The Journal of APPLIED PSYCHOLOGY

Edited by

JAMES P. PORTER
Ohio University

WILLIAM F. BOOK
Indiana University

And a Board of Co-operating Editors

VOLUME VI

1922

Published by the Editors
INDIANA UNIVERSITY PRESS
BLOOMINGTON, INDIANA

311348

Copyright, 1922, by the Editors.

**INDIANA UNIVERSITY PRESS
BLOOMINGTON, INDIANA**

TABLE OF CONTENTS

ROSE G. ANDERSON	
Methods and results of mental surveys . . .	1- 28
ADA HART ARLITT	
Relation of Intelligence to age in negro children .	378-384
HARRY J. BAKER	
Mental tests as an aid in the analysis of mental constitution	349-377
F. S. BREED AND S. E. KATZ	
The color preferences of children	255-266
R. A. BROTEMARKLE	
A comparison test for investigating the ideational contest of the moral concepts	235-242
AGNES BURKE, LETA S. HOLLINGWORTH AND CHARLOTTE G. GARRISON	
Subsequent history of E———; five years after the initial report	205-210
J. CROSBY CHAPMAN .	
The individual injustice and guessing in the true false examination	342-348
RUTH SWAN CLARK	
A glimpse of high school courses as measured by the Otis test	185-191
KARL M. COWDERY	
Measures of general intelligence as indices of suc- cess in trade learning	311-330
JUNE E. DOWNEY AND LOUISA C. WAGONER	
Speech and will-temperament	291-297
MAX FREYD	
A method for the study of vocational interests .	243-254
CHARLOTTE G. GARRISON, LETA S. HOLLINGWORTH AND AGNES BURKE	
Subsequent history of E———; five years after the initial report	205-210
THOMAS R. GARTH	
Mental fatigue of mixed and full blood Indians .	331-341
DOROTHY B. GOLDSMITH	
The use of the personal history blank as a sales- manship test	149-155
AMY HEWES	
Standardization of the Whipple-Healy Tapping Test	113-119
LETA S. HOLLINGWORTH, CHARLOTTE G. GARRISON AND AGNES BURKE	
Subsequent history of E———; five years after the initial report	205-210

CONTENTS

CLARK L. HULL

- The conversion of test scores into series which shall
have any assigned mean and degree of dispersion 298-300
- The computation of Pearson's r from ranked data 385-390

EDWARD S. JONES

- Effect of letters and syllables in publicity . . 198-204

O. J. JOHNSON

- Group Intelligence examination for primary pupils 403-416

S. E. KATZ AND F. S. BREED

- The color preferences of children 255-266

HARRY DEXTER KITSON

- Minor studies in the psychology of advertising . 59- 68

V. The development of art-forms in magazine
advertising

VI. Color in magazine advertising

VII. Negative suggestions in advertising

- A shift of emphasis needed in personnel research 141-148

✓ F. B. KNIGHT

- The significance of unwillingness to be tested . 211-212

R. L. LEWY-GUINZBURG AND SANTE NACCARATI

- Hormones and Intelligence 211-234

FRANCES LOWELL

- An experiment in classifying primary grade child-
ren by mental age 276-290

SANTE NACCARATI AND R. L. LEWY-GUINZBURG

- Hormones and Intelligence 221-234

— A. T. POFFENBERGER

- The selection of a successful secretary . . . 156-160

A. T. POFFENBERGER AND V. H. VARTANIAN

- The letter of application in vocational selection . 74- 80

M. J. REAM

- A social relations test 69- 73

MRS. ANNE ROSS REYMERT

- Some factors of aesthetic judgment 34- 58
120-140

W. T. ROOT

- The intelligence quotient from two viewpoints . 267-275

JOHN SLAWSON

- The reliability of judgment of personal traits . 161-171

EDWARD L. THORNDIKE

- On finding equivalent scores in tests of intelligence 29- 33

HERBERT A. TOOPS

- Computing Intercorrelations of tests on the adding
machine 172-184

CONTENTS

CLARA HARRISON TOWN

A mass mental test for use with kindergarten and
first grade children 89-112

EGBERT M. TURNER

The testimonial as an advertising appeal . . . 192-197

V. H. VARTANIAN AND A. T. POFFENBERGER

The letter of application in vocational selection . . 74- 80

MORRIS S. VITELES

A comparison of three tests of general intelligence 391-402

LOUISA C. WAGONER AND JUNE E. DOWNEY

Speech and will-temperament 291-297

NOTES 81-82, 211-214, 301-302, 417-428

BOOK REVIEWS 83-88, 215-220, 303-310, 429-432



6. Athena and the Olive Branch



1

2

3

4

5

6

7

8

9

3 6 2-

The Journal of
APPLIED PSYCHOLOGY

Edited by

JAMES P. PORTER
Ohio University

WILLIAM F. BOOK
Indiana University

And a Board of Co-operating Editors

VOLUME VI

1922

Published by the Editors
INDIANA UNIVERSITY PRESS
BLOOMINGTON, INDIANA

The purpose of this survey was to determine the number of mental defectives of institutional grade in the county. A study was also made of children of superior intelligence. S. L. Pressey and L. W. Pressey had charge of the mental testing done in this survey and devised a group scale of intelligence which was given to all school children in grades 3-8 inclusive in the schools surveyed, as a means of selecting those cases which needed further individual examinations. A considerable number were then examined individually with the Stanford Revision. The schools surveyed were the rural and village schools in nine out of the twelve townships in the county, and the schools of Bloomington. The average percent of feeble-mindedness for the schools outside of the county seat was found to be 6.4, averages for individual townships varying from 2 percent to 20 percent. In the schools of Bloomington 2 percent were feeble-minded.

DISCUSSION OF SURVEYS

It will be seen that these surveys show a wide variation in the proportion of feeble-mindedness in the population.¹¹ The smallest percent appears in a study which was not primarily an effort to discover all cases of feeble-mindedness, but to learn under what conditions the known cases live, and the largest percent, in a study where children who were likely to be cases of mental deficiency were selected by the use of group tests and teachers' estimates and were given a careful individual examination with standardized mental tests. In most instances, the smallest ratios found are in surveys in which information in regard to the numbers of cases was secured from individuals and agencies not especially qualified to diagnose mental deficiency. The more careful and intelligent these observers were, the more certain they would be to hesitate to classify as feeble-minded, any cases which were at all doubtful, thus missing the greater number of higher-grade cases which to the untrained observer so frequently appear normal.

It is an oft-repeated experience that well-intentioned individuals prefer to excuse indications of mental inferiority in others on the basis of ill-health, poor home care, language difficulty, lack of opportunity, etc., unless the case is a troublesome or delinquent one. This is especially true in the case of teachers' judgments even in cases where a child is several

¹¹ See Miner, *op. cit.*, for reviews of other surveys. Also Mental Hygiene, Vol. 1, No. 1, on a report of the survey of a section of Baltimore by Dr. C. Macfie Campbell.

years retarded pedagogically. The age-grade status of the child, the most important criterion at the teacher's disposal, is the most frequently disregarded. Since, as is well known, these higher grade feeble-minded constitute the majority of the total number of feeble-minded, their omission would materially reduce the percent of feeble-minded. The writer does not doubt, that the various areas surveyed differ somewhat in their respective percentages of mental deficiency, but it is also her belief that most of these differences can be explained by the difference in the methods of conducting the surveys and in the standards used by the various investigators in diagnosing mental deficiency. Further discussion of the surveys should serve to bring this out.

In the Ohio survey, in the case of school children, all children who were three years or more retarded pedagogically, together with others named by the teachers, were chosen for individual consideration and the effort was made to predict from the known facts, whether the child would be able to maintain an independent existence when mature. Considering the difficulty of making such a prediction with no fixed standards of comparison, it is surprising that as large a percentage as 1.8 was classified as definitely feeble-minded. But, on the other hand, many children, especially younger children, who are less than three years retarded pedagogically, prove to be definitely feeble-minded. Mental retardation is present in a defective child when he enters school whether he is entered by his parents at the normal age or not. A defective child would have to remain in the first grade three years before he would be considered a suspect if the above basis was used for selection, whereas, at the age of six he was practically as mentally retarded as at the age of nine. It is true that some children were finally classified as feeble-minded where the pedagogical retardation was less than three years, but the proportion is very small and from the description, the cases were evidently very low-grade, at least to be attending public school. So that it seems safe to conclude that many higher grade cases of defect must have been overlooked with the method of selection used.

It is often stated that there is a marked tendency on the part of the lay public not to report a case as feeble-minded unless it is quite low-grade or unless it is anti-social in its reactions. In the New Castle County survey by the Children's Bureau, 1100 cases were reported to the investigators as possibly feeble-minded. Only about a fifth of this number was classified as positive cases and an additional one-third

as questionable cases. It is stated that a number of cases was dropped because of the difficulty of diagnosis or because of lack of sufficient evidence to classify them as defective. In the report of the school cases which were defective, the least retarded cases in the various grades have been noted above. It is evident from an examination of these, that many cases of milder defect have undoubtedly not been included. Miss Lundberg clearly recognized that the numbers quoted did not include all the mentally defective cases and that they could only represent the minimum of all cases in the county.

The New Castle County survey by the United States Public Health Service confined itself entirely to the school children of the county and, as has been noted, gave a considerably larger percent of mentally defective and probably mentally defective children. In a number of the surveys reviewed, it is evident that there must have been a good deal of variation in the judgments of different investigators engaged in the same survey. Although the fact is not stated, it is evident that a number of investigators must have been engaged simultaneously in this survey as 3793 children were studied in four months. A part of the preliminary examination given consisted of "questions suited in a general way to the child's age and school grade," and a part of the secondary examination consisted of tests and exercises that could not help but vary a great deal in the case of different examiners when these tests and questions had not been previously standardized or were incapable of standardization. Even though the percentage quoted here is higher than some of the others, it is evident that it is too low when cases which have a chronological age of 15 and a mental age of 10.6, or a chronological age of 14 and a mental age of 9.8 are accepted as normal because so considered, in spite of the above evidence, and even though the fact that the Goddard Scale measures somewhat short for these ages is taken into consideration.

A consideration of the basis used for classifying children as exceptionally retarded in the survey of Porter County, Indiana, will suffice to show the possibility of overlooking cases of defect. Although perhaps not exactly comparable for the Goddard Revision, it will bring out the extent to which the levels at which a child was considered exceptionally retarded varied for the different ages, if the mental and chronological ages quoted are changed into intelligence quotients for each age. A six year child with 67 percent intelligence was so classified, a seven year child with 71 percent, an eight year old with 63-75 percent, a nine year old with 67-78 percent, a

ten year old with 60-70 percent, an eleven year old with 64-73 percent, a twelve year old with 67-75 percent, a thirteen year old with 61-69 percent, a fourteen year old with 64-71 percent, a fifteen year old with 60-67 percent, a sixteen year old with 67 percent, and a seventeen to a twenty year old with 73 percent, when fifteen years is used as the divisor for the ages above fifteen. Further standardization of the Binet-Simon tests and of the earlier revisions has shown that the tests for the lower ages are too easy so that consequently the ratings of the younger children by the Goddard Revision are too high. Without considering this fact, it is evident from the above percents, that it is highly probable that numbers of cases of feeble-mindedness were not so classified, and that there was much possibility for variation in judgments especially for those ages, 8-15, for which two mental ages are given.

The Maine survey did not attempt an enumeration of all cases, and only studied a very small proportion individually, which is sufficient to account for the small percentage of feeble-mindedness which the investigators found.

There are several outstanding features of the survey of "X" County, California. The first of these is the careful selection of cases by the use of three criteria, (1) teachers' judgments of school work, (2) teachers' estimates of ability, and (3) the amount of pedagogical retardation; second, it would seem that a large enough proportion was examined individually to include all suspected cases; third, these cases were examined by a carefully standardized scale and by investigators whose preparation was directly in this field, and who had had sufficient training in this field to make their results authoritative, especially when verified by Dr. Terman. Some of the other surveys have had one or more of these features, but the writer believes none of them has fulfilled so well the requirements of a well-conducted survey. Dr. Terman considers the percent of feeble-mindedness found in this study "astonishingly high," which it is when compared with the results of others, and concludes that "the ratio in 'X' County indicates an exceptional condition." The writer would not venture to deny this conclusion, but would again suggest that the disparity in the results of this and other surveys is due, in large part, to the differences in the methods of conducting them.

The Presseys in the survey of "X" County, Indiana, went one step further in improving upon the method of making a mental survey, by introducing the use of group tests. By their use an approximate mental rating was obtained of every

child in grades 3-8 inclusive in all the schools surveyed, thus making the chances even less of not detecting all cases of feeble-mindedness. Group tests were not given, however, in grades 1 and 2. It is in these lower grades that the probability of error in the teachers' selections is greatest. The tendency of the schools is to keep the younger children under observation for several years, before reporting them as mentally inferior, and then to report a child who is noticeably over-age only in case he is also doing inferior work. In many instances, feeble-minded children in the first and second grades are not reported to the examiner because they have not been in school long enough to become sufficiently retarded to be regarded as likely cases of mental deficiency. In other cases, where they are sufficiently over-age for their grade so that they are able to do passing work, this latter consideration is sufficient to exclude them from a list of supposed mental defectives. For the greatest reliability in results, it is necessary to have a means of checking the teachers' selections from these grades as well as from the upper grades.

METHODS AND SCOPE OF MINNESOTA SURVEYS

Simultaneously with the beginnings of the survey of "X" County, Indiana, the Research Department of the Minnesota School for Feeble-Minded began a survey of "X" County, Minnesota. One of the first steps in the preparations for making this survey was the devising of a group test scale by Miss Frances Lowell who began the work of the survey.¹² This scale was revised slightly by Miss Maude Merrill and during the past three years has been further revised and extended by the writer, all of the work on the scale being done under the general direction of Dr. F. Kuhlmann. At present the scale contains groups of tests applicable to each age from five years to maturity. As the survey is not yet completed, nor the group tests ready for publication, it is not intended here to give the results of the county survey as a whole nor to present the group test scale, but to present the methods and results of the surveys of: (1) a group of 23 city school systems in all parts of Minnesota having grade populations varying from 162 to 2640; (2) a group of three city school systems, "A," "B," and "C," having grade enrollments of 962, 1301, and 1490, respectively; and (3) a group of 35 rural and small village schools with a total enrollment of 757. In the first group, selections of cases for individual

¹² Lowell, Frances, A Group Intelligence Scale for Primary Grades, *Journal of Applied Psychology*, Sept., 1919, Vol. III, pp. 215-247.

examinations have been made by the teachers in their respective schools; in the last two groups, the method of selection for individual examinations has been by the use of the group tests. In the discussion, these two methods of selection will be referred to as the 'group-test-selection' and 'teacher-selection' methods. The study comprises a total of 1490 individual examinations and 3524 group test examinations.

Minnesota provides state-aid for special classes for sub-normal children in the public schools throughout the state. These classes are limited to an enrollment of fifteen and to children having an intelligence quotient of 50-85. Examinations of candidates for these classes are made by the Research Department of which the writer is a member. The method of making a survey of a school system to determine those eligible to the special class consists first of selection by the teachers of all cases in their own rooms which they consider sufficiently retarded to be eligible to the special class. Preliminary information blanks, supplied by the state supervisor of special classes, which call for the complete school record of the child, information in regard to the parents, home care the child receives, etc., are then prepared. In addition to this, each child is given a medical examination. These data are prepared before the mental examiner arrives and are made available before the mental examination takes place. The plan followed is to test all cases selected by the teachers whether the class or classes can accommodate them or not. In the opinion of each school system, the survey has included all cases suspected of mental deficiency and is complete. The examinations in the group of 23 schools mentioned above have been made in connection with such surveys for special classes. They represent a selected group of schools which have been surveyed in this manner. In recent years group tests have been given in a number of places as an aid in selecting the cases for individual examination; in others, the special class teacher, who has often taken a short course in testing, has examined the children in order to have a more selected list for the state examiner. Since such cases would not of course be a fair sample of the 'teacher-selection' method, only those schools where no preliminary mental testing has been done, and where examinations were made for the first time have been included in the group under consideration.

City "A" and all the schools in group 3 are in "X" County and have been surveyed as a part of the county survey. City "A" is the largest city and the county-seat of the county. The group of 35 rural schools is not a selected group except in

that these schools are the most recently surveyed of the rural schools of the county. Cities "B" and "C" of group 2 are representative cities in other parts of the state which were surveyed, (1) as a means of comparison with city "A," and (2) as a means of comparison of the 'group-test-selection' method with the 'teacher-selection' method.

All examinations, both group and individual, have been made by the members of this research department. The Kuhlmann Revision of the Binet-Simon tests has been used for individual examinations.¹³ As reliable results can be obtained through the use of this revision as are at the present time possible, because of the thoroughness and accuracy of its standardization, the large number of tests for each year from 3-15, and its continued improvement through its several revisions. An advantage has been that an abbreviated scale consisting of largely 'non-verbal' tests has been derived from it for use on cases who suffer from a language handicap or who, because of excessive timidity, would not respond as readily to the tests which require longer verbal responses or the formulation of answers. Comparison of this 'half-scale' examination with the full-scale examination of a large number of cases of the above description showed that the score obtained on the 'non-verbal' tests alone was always higher than that on the entire scale. In a considerable number of cases in the rural schools, this 'non-verbal' examination was used in place of the complete examination. It should be borne in mind in the consideration of the rural school results that the possibility of error in these test results is in the direction of being too high rather than too low, which would have been the reverse had the full-scale examination been given to these cases.

As the medical and social data necessary to the final diagnosis of borderline cases have not been completed in the case of the rural school examinations and those of cities "A," "B," and "C," the cases examined have been classified on the basis of the individual mental examination results alone. Cases with an I. Q. of 50-74 are classified as morons; 75-84, as borderline; 85-94, as dull; 95-104, as average normal; 105-114, as bright, etc. This classification is of course somewhat arbitrary and it is possible that some cases are included in the moron group, who in the light of the completed data would not necessarily be classed as feeble-minded in the narrower sense,¹⁴ but on the other hand, it is certain that

¹³ Kuhlmann, F., *Handbook of Mental Tests*, Warwick & York, Baltimore, 1921.

¹⁴ That is, requiring institutional care or supervision because of failure, due to lack of intelligence, to pass the social test.

many cases testing 75 or a few points above would, with extended information, be classed with the definitely feeble-minded. It is possible that these two changes might compensate each other. A recent extensive study by Dr. Kuhlmann has shown that there is a general tendency for the intelligence quotients of defective children to decrease as they grow older.¹⁵ Those cases who in childhood have an I. Q. of 75 or a few points above or below will in all probability not maintain this I. Q., and by the time they are adults, it will be somewhat lower. Dr. Kuhlmann has concluded that the chances are practically nil "that such adults (below an I. Q. of 75 or 70) will be able permanently to make an independent, honest living, without supervision and guardianship, under any and all circumstances they are likely to meet in their lives."¹⁶ Other considerations aside, all those with an I. Q. of 50-74 are as mentally inferior as the high grade feeble-minded. They constitute the group which is the greatest problem to the public school and society, and to which school and society are problems of like magnitude. It is of importance to know how numerous this class is, in order to have a more comprehensive program to deal with the difficulties incident to its members. To attempt to minimize the importance of the problem, by presenting data which are only a partial picture of the conditions, can only have results disastrous to the efforts of those attempting to deal with conditions as they actually are.

COMPARISON OF TWO METHODS OF SELECTION IN CITY SCHOOLS

TABLE I

	Total No. Cases	No. Indiv. Tests	Percent. Indiv. Tests	F. M.	Border Line	Dull	Av. or Above	Percent of Total F. M.
'Teacher— selection'	20,877	911	4.4%	50%	25%	18%	7%	2.2
'Group-test— selection'	3,001	121	4.4%	67%	25%	8%	0%	3.1

Of the total enrollment in the 23 cities 911 cases or 4.4 percent were selected by the teachers for individual examinations. In the three cities surveyed by the group test method, the proportion of individual examinations was considerably greater. The total number examined in the latter was 315 or 10.5 percent of the number given the group test. In order

¹⁵ Kuhlmann, F., The Results of Repeated Re-Examinations of 639 Feeble-Minded Over a Period of Ten Years, *Journal of Applied Psychology*, Sept., 1921, Vol. V, pp. 195-224.

¹⁶ Kuhlmann, F., Determination of Feeble-Mindedness as Related to the Courts, Minnesota State Board of Control, Saint Paul, 1920.

to have the results of the two types of surveys more nearly comparable, the individual ratings of the 4.4 percent who received the lowest ratings in the group tests are compared (Table I) with the 4.4 percent chosen by the teachers in the 23 cities. In the latter group, 75 percent rated feeble-minded or borderline in the individual test, and the remaining 25 percent rated dull, average, or slightly above the average. The cases rating feeble-minded constituted 2.2 percent of the total school enrollment. Out of the 4.4 percent lowest in the group tests, 92 percent rated feeble-minded or borderline in the individual test and 8 percent dull, with none rating above this. This lowest 4.4 percent in the group tests yields a percentage of 3.1 percent of the total rating feeble-minded. Out of a hundred children, selected by each of the two methods, for individual examinations the group-test-selection method discovers 17 more cases of mental deficiency, and eliminates, on the other hand, 17 individual examinations of cases too high to classify as either feeble-minded or borderline.

The results of the completed survey show that the percentage considered above is not large enough to reveal all cases below an I. Q. of 75. As has been stated, 10.5 percent of the total number reached in the group test were also given individual tests. It is obvious that the larger the percentage tested, the greater becomes the possibility of error in selecting the cases, because the scores approach more nearly the average and a smaller gain in the individual test will put the case into the dull or average group. In spite of this fact, the larger percentage of 10.5 percent compares favorably with the 4.4 percent selected by the teachers. Out of the total number of 315 cases examined individually, 77 percent classified as feeble-minded or borderline, and 23 percent as dull or average, compared with 75 percent and 25 percent, respectively, for the other method. The percent of cases rating below an I. Q. of 75 were: City "A"—4.5 percent; City "B"—3.5 percent; and City "C"—4 percent, an average of 4 percent for the three taken together. The percents quoted for cities "A" and "C" are low because the former system already had one special class and the latter, two special classes organized before the group test survey. The cases in these classes are not included in these results. The Presseys found that all the cases of mental deficiency were included in the lowest 10 percent in the group test. This agrees very closely with the percent it was found necessary to examine individually in the above schools. With further improvement in the group tests it may be possible to reduce this percentage somewhat.

The question arises as to how representative these results are likely to be of the population as a whole. Any study, to be representative, must include a large group of unselected cases. The only place where it is at all possible to find such a group is in the public schools, especially in the lower grades. Even in these selection has begun to occur because low grade cases have not been entered at all or have been excluded shortly after their entrance. The percent of these is small, however, and their omission does not affect the results appreciably. In the upper grades, elimination takes place, the lower cases here dropping out after they have passed the age for compulsory school attendance. But, on the whole, the sampling secured here is the most representative that is possible, and the distribution of grades of intelligence is undoubtedly approximately the same as would be found in any equally representative group at any age. As Miner has indicated,¹⁷ the fact that mental deficiency is not a condition overcome with age, but a permanent state, assures closer agreement between the percent of feeble-minded in the school population and the general population than most investigators have been accustomed to quote, even when allowance is made for the greater death rate among the feeble-minded.

At the time of its use in these surveys, the group test scale was by no means as accurate as it will be when finally standardized. When the scale is in its final state, it is expected that the efficiency of the group-test-selection method over the teacher-selection method will be even greater. Because of the individual variations which are always present in the reactions of certain types and ages of children in a group test and an individual test, there will always be a certain amount of variation between the results secured by the two types of scales, however carefully standardized, which it will be impossible to control. For this reason the group test can never hope to supplant the individual test nor give a final judgment on a case.

This added efficiency in the selection of cases is only one of the advantages of this type of survey. The great saving in time and expense must not be overlooked. The more accurate and efficient selection by the group-test-selection method has been possible after only a fraction of a day spent in each schoolroom, whereas, in the case of the teachers' selection, it has usually been necessary for a child to repeat one or more grades before the evidence is considered sufficient to place him in the suspected group. Individual mental test results

¹⁷ *Op. cit.*

of a number of children ranking highest in the group tests indicate that the method is just as efficient for the selection of the most superior children for further study.¹⁸ Comparisons of the averages for different buildings and grades are possible and often reveal unsuspected and startling discrepancies between the average levels of ability of the same grade in different buildings or of different grades in the same building. The problem of the teacher who has, on the average, less well-endowed pupils is better understood, and the failure of these pupils to measure up to others in pedagogical tests need not then reflect discredit on the conscientious teacher, as is otherwise apt to be the case. Cases have been reported of certain teachers receiving additional recognition because of the rank of their pupils in pedagogical tests. Intelligence tests of these same pupils have shown that, in reality, their attainment was below what it should have been when their native ability was taken into consideration.

DIFFERENCES IN GRADE AND BUILDING AVERAGES

It was expected that the average levels of ability in the various grades and buildings would differ somewhat. In certain instances, however, the contrasts between grades were so pronounced that a study was made of the results for the three cities "A," "B," and "C" to determine how general and how extended such differences were. The results of this study are so suggestive that it was thought advisable to include them in this report. The variation in average ability of the different grades and buildings and the variation in the distribution of the brightest and dullest children of each system among the different buildings will be shown.

With the group test scale that has been used, it has been possible to secure a mental age and an I. Q. for each child taking the group test. A comparison of the results of the individual test and the group test for the same children has shown that the group test measures slightly short, that is, the mental age secured by the group test is, in most instances, lower than that secured by the individual test. In order to have the average group test rating more nearly correct, the group test scores have been corrected, so that the average score for each chronological age is 100, before the averages for the grades and buildings have been secured. This assures the accuracy on the average of the group test results. As a check on the accuracy, the average 'corrected' group test

¹⁸ See Book, *op. cit.*, for data on this point.

score was compared with the average individual test score for all the children in one of the cities ("B") who were given individual tests. The former score was found to be .3 of one point higher than the latter, so that the averages can be accepted as reliable estimates of the average abilities of the various grades and buildings.

TABLE II
AVERAGE SCORES FOR GRADES AND BUILDINGS

Schools	Grade I	Grade II	Grade III	Grade IV	Grade V	Grade VI	Bldg.	City
City "A"—I	107	103	104	101	101	105	104	
II	98	107	100	103	105	107	103	
III	90	97	99	107	107	104	101	
IV	95	93	96	95	89	105	96	
V	85	99	93	98	94	94	94	
								100
City "B"—I	96	106	102	100	103	104	102	
II	95	98	101	99	101	104	100	
III	93	97	98	96	94	100	96	
IV	97	95	94	97	94		95	
								98
City "C"—I	98	104	102	101	107	97		
			96				101	
II	104	102	96	97	97	101		
	92			93	96	92		
					92		97	
III	88	99	101					
		94					96	
IV	96	97	95	94	96	94	95	
V	93	101	97	102	96	93		
				88	93	87	94	
								97

TABLE III
AMOUNT OF VARIATION IN SAME GRADE IN DIFFERENT BUILDINGS

	Grade I	Grade II	Grade III	Grade IV	Grade V	Grade VI
City "A"	22	14	11	12	16	13
" "B"	4	11	8	4	9	4
" "C"	16	10	7	14	15	14

TABLE IV
AMOUNT OF VARIATION IN DIFFERENT GRADES IN THE SAME BUILDING

	School I	School II	School III	School IV	School V
City "A"	6	9	17	16	14
" "B"	10	9	7	3	
" "C"	11	12	13	3	15

In Table II are given the average scores for each grade, building, and city. From Table II, Tables III and IV have been derived, summarizing the range in scores in the different grades and buildings. In City "A" there is a variation of 22 points in the average scores of the first grades in the different buildings. The significance of this difference is more clearly seen if the difference in I. Q. is translated into the corresponding difference in mental age. If the assumption is made that the average chronological age in the first grade was seven years,¹⁹ the average mental age for Grade I in School V was 5.95, and for Grade I in School I was 7.49, an average difference of 1.54 years in mental age between two grades in the same system from which the same results are expected. Assuming the average chronological age for the second grade to be eight years, for the third, nine years, etc., the same comparisons can be made for the other grades. Perhaps the large differences between the same grade in different buildings are not so unexpected. It is more surprising to find that almost as large and as frequent discrepancies occur in the different grades in the same building. Tables (II-IV) should serve to bring out forcibly the fact that it is not sufficient to have a general notion that one grade is better or poorer than another, but that it is necessary, for the best results, to have an accurate measure of how much above or below average a grade or building is.

DISTRIBUTION OF EXTREMES IN ABILITY AMONG GRADES AND BUILDINGS

As a further means of comparison of the different systems, the highest 10 percent and the lowest 10 percent in ability were selected in each system, and the distribution of these best and poorest among the different buildings was studied. These results are shown in Plates I and II. In Plate I, the percent of each building's enrollment which is in the 10 percent lowest and 10 percent highest of the cities is shown. If these were distributed uniformly each building would have 10 percent of its enrollment in each of these groups. However, the distribution is far from uniform. In city "A," School IV has four times as large a percent of its enrollment among the 10 percent lowest than School II, and on the other hand, School I has over three times as large a percent of its enrollment in the 10 percent highest than Schools IV and V have. The contrasts are almost as striking in cities "B" and "C."

¹⁹ All of the surveys were made at or after the middle of the school year.

PLATE I

PER CENT OF EACH BUILDING'S ENROLLMENT IN THE
10% LOWEST AND HIGHEST OF THE CITY

	LOWEST	RANGE IN GRADE AV.	AV. SCORE	HIGHEST
CITY "A" - SCH. I	9% [bar]	101-107 [bar]	104 [bar]	17% [bar]
" II	4% [bar]	98-107 [bar]	103 [bar]	12% [bar]
" III	6% [bar]	90-107 [bar]	101 [bar]	14% [bar]
" IV	16% [bar]	89-105 [bar]	96 [bar]	5% [bar]
" V	14% [bar]	85-99 [bar]	94 [bar]	5% [bar]
CITY "B" - SCH. I	5% [bar]	96-106 [bar]	102 [bar]	17% [bar]
" II	7% [bar]	95-104 [bar]	100 [bar]	11% [bar]
" III	16% [bar]	93-100 [bar]	96 [bar]	4% [bar]
" IV	12% [bar]	94-97 [bar]	95 [bar]	6% [bar]
CITY "C" - SCH. I	7% [bar]	96-107 [bar]	101 [bar]	16% [bar]
" II	10% [bar]	92-104 [bar]	97 [bar]	10% [bar]
" III	5% [bar]	98-101 [bar]	96 [bar]	4% [bar]
" IV	7% [bar]	94-97 [bar]	95 [bar]	8% [bar]
" V	14% [bar]	87-102 [bar]	94 [bar]	9% [bar]

PLATE II

DISTRIBUTION OF LOWEST AND HIGHEST 10% OF
EACH CITY AMONG THE DIFFERENT BUILDINGS

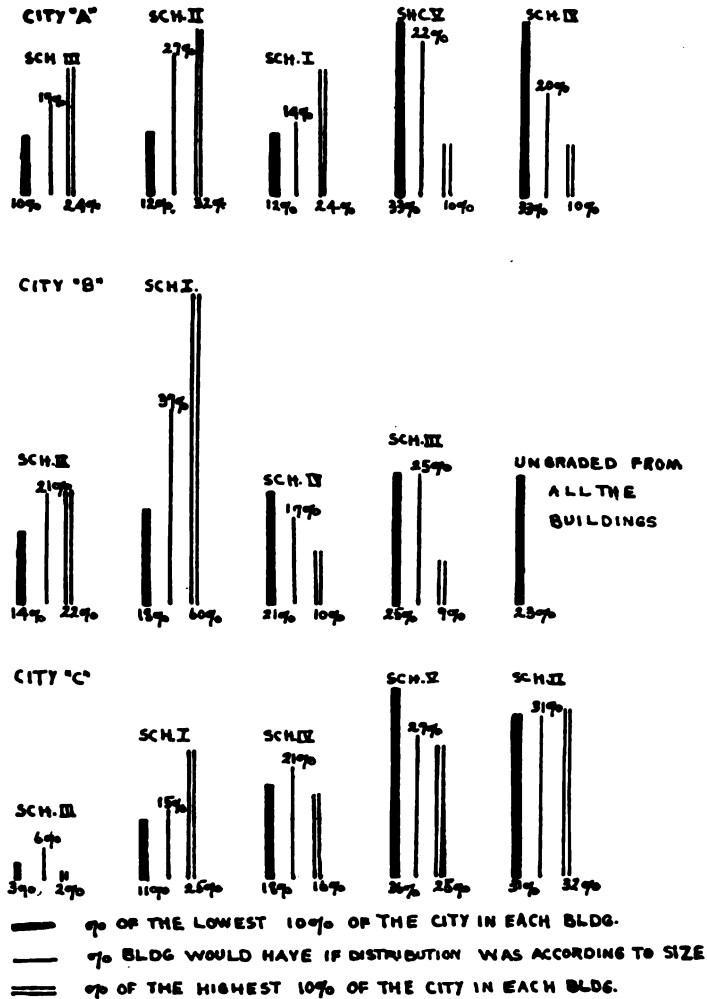


Plate II gives a comparison from another point of view. The distribution among the buildings of the children ranking in the best and poorest groups is here compared with the percent the building would have if the best and poorest were distributed in proportion to the size of the school. In general the schools which have the highest average scores have less than their proportion of weak students and more than their proportion of strong students, the reverse being also true.

As would be expected the schools in the better districts rate higher than the schools in the poorer districts. The agreement between the rank of the building and the type of neighborhood in which the building is found is very close. In every case, the schools which rate the lowest and have the larger proportion of dull children and the smaller proportion of bright children are situated in the part of the city where the homes are poorer and where many of the fathers are unskilled laborers and are unable to provide homes of the better sort. The schools ranking the highest and having the smaller proportion of dull children and the larger proportion of bright children are situated in the part of the city in which the professional and business men of the city live and in which the best homes are found. This fact often gives rise to the erroneous belief that the children from the poorer homes rate lower because of the lack of advantages in their environment, when, as a matter of fact, the poor environment and lack of advantages are in most instances the accompaniments and results of the lower mentality of the parents instead of the cause of the lower mentality of the children. The contrasts in the different buildings would undoubtedly be much greater in larger cities where the better residence districts and slum districts are more differentiated.

RURAL SCHOOL SURVEYS

Although the percentage of feeble-mindedness in the city schools is higher than would be expected by many, it seems low when compared to the percentage found in the rural schools. Because of the large number of cases of feeble-mindedness found in the latter, the writer must confess some hesitancy in presenting the rural school results.

The work in these schools was done in the fall when the weather was still sufficiently warm to permit the use of the small entrance hall to the school-house for the individual examinations, thus avoiding disturbances, interruptions, and lack of response on the part of the subject because of the presence of others. The examiner arrived at about the time

school started in the morning and the group test was given immediately, everyone taking it at once. The tests for the youngest age groups were given first and the explanation was made that some easier things would be given first so that the younger children could do them too. As the younger children automatically dropped out because of the increasing difficulty of the tests, their books were taken up and the teacher gave them seat work to do. As this procedure lengthened the process of giving the group test considerably, a recess was given when the upper age tests were reached. During this time the examiner looked hastily through the books and took up those in which the subject had obviously gone beyond his limit. The rest of the tests were then given and the remainder of the morning was utilized for scoring the results. The children obviously enjoyed the tests and were stimulated to their best effort when they were told that the same ones were being given in all the schools of the county and that one of the objects was to see which school could do them the best. At noon the examiner ate her lunch with the teacher and children and often joined in their games, so that by the time school was called in the afternoon any feeling of timidity had worn off. The cases chosen for individual examinations were plainly envied because of this distinction. The explanation was made that only a limited number could be taken because of the time factor. Where ever there was an indication of language difficulty (some of the schools were in Bohemian districts) or of hesitancy in responding to the verbal type of tests, the 'non-verbal' examination already referred to was given. It would seem that conditions were as favorable as those usually obtainable. In smaller schools where only one or two individual examinations were necessary, it was possible to finish the school in a day, in larger ones, it was necessary to return for a second day or fraction thereof. Only one school had more than one teacher. This one had three teachers and was given three days time.

In the 35 schools, individual examinations were given to 35 percent, 183 cases, of the total number taking the group tests. Five of these schools had no cases with an I. Q. below 75, nine had from 1-10 percent, 13 had 10-25 percent, and eight had over 25 percent of those present rating below an I. Q. of 75, an average of 15 percent for all the schools taken together. Besides this an additional 11 percent rated between 75 and 84 in the individual examination. In two schools no individual examinations were necessary, in fifteen schools more than 35 percent of those present were examined, nine of these

having over 50 percent of those present examined. Eleven of the schools had 66 percent of all the cases below an I. Q. of 75.²⁰ This large percentage of individual examinations was necessary because of two things; the large number of cases testing below 75 and a few points above, and the fact that the rural school children are more accustomed to being dealt with individually or in very small groups, and therefore their performance in the larger group was often a good deal poorer than it was in the individual test. This was especially true of the younger children.

The large percentage of cases below an I. Q. of 75 may signify one of several things. It may show, as Pressey has asserted, that the standardized intelligence scales in use have features that are distinctly urban in nature and that rural school children should not be judged by them on the same basis as city children. Another conclusion possible is that these schools are not representative of rural schools as a whole, but that their extremely high percentage of cases of low mentality represents an abnormal condition and is disproportionately high for rural communities in general. On the other hand, it may serve as conclusive proof of the generally-held opinion that a higher grade of mentality is required to maintain an existence in an urban community than is required in a rural community.

In regard to the first, an examination of the tests in the scale used reveals only a very few tests which might be passed with greater ease by a city child than by a rural child of the same mentality. But even though this were true, it is possible that tests which might serve as a truer index of a rural child's mental capacity could be substituted for some of these, and that the lower rating of the rural children may be due, in part, to

²⁰ It has been noted that the average percent of feeble-mindedness for the townships in the survey of X County, Indiana, varied from two to twenty. In these 35 rural schools in Minnesota, the variation was from 0 to 50 percent for schools, and from 6 to 20 percent for townships. These large discrepancies in the results within the same survey might seem to weaken the writer's contention that the differences in the findings of the surveys discussed were due largely to the differences in the methods used in conducting the surveys and in diagnosing mental deficiency. The writer has not attempted to show that these differences were always or entirely due to these differences in methods. Local differences are granted and expected. These large differences in neighboring schools and townships emphasize the necessity of obtaining a sampling from a large enough area to represent an average of conditions. All of the surveys discussed have fulfilled this condition. It is where averages from such representative samplings vary from .212 percent, .324 percent, and .44 percent to several times this proportion that the question of method is pertinent.

the omission of certain types of tests rather than to the unfairness of tests already included. At first thought it would seem that this could only account for a small percentage of the difference between the city school and rural school results. However, closer inspection reveals that this factor might influence the percent rating mentally deficient very materially. Of the total number rating below an I. Q. of 75, 38 percent fall in the range 70-74; 53 percent in the range 60-69; 7.5 percent in the range 50-59; and 2.5 percent below 50. Suppose the assumption is granted that the tests used measure slightly low for rural school children, and that the limit for arbitrary classification of cases as feeble-minded is dropped five points. Those cases falling below an I. Q. of 70 would then be classified as feeble-minded. This drop of five points changes the percent of feeble-minded from 15 to 9.5, a very considerable reduction. The need of further research along this line is suggested. It is not improbable that further study might prove that the above assumption is justifiable.

The county being surveyed contains fourteen townships. Seven of these are represented in the group of schools under consideration, and only two of these by less than five schools. The results of the completed survey may show that conditions in the other townships are better than in those represented in this study. It is certain that one of the townships, which contains six of the schools out of the 35, is not inhabited by enterprising farmers because of the poor quality and unfavorable nature of much of the land in it. The percents of cases below an I. Q. of 75 in these schools are: 7, 10, 22, 36, 40, and 50. This last school had an additional 30 percent which graded as borderline cases. The enrollment of the school was made up of children from five families. The children of two of these, which comprised all but four of the total enrollment, were cousins. One of the remaining four was a middle-grade imbecile and was almost totally blind, one was a borderline case, and the remaining two were a brother and a sister who were normal. Even if the majority of the cases of mental deficiency in the county were in these 35 schools, the percentage for the county would still be extremely high compared to that quoted by other investigators. It is probable that the county as a whole is not representative of Minnesota.

That this is undoubtedly not a representative Minnesota county is indicated by the 1920 United States census report for Minnesota.²¹ This report includes figures on the percent

²¹ United States Bureau of the Census, Fourteenth Census of the United States: 1920, Population: Minnesota, Composition and Characteristics of the Population.

of illiteracy in the counties and cities of Minnesota. For the state taken as a whole the percent of illiteracy is 1.8. Sixty out of the 86 counties have less than two percent, 21 have two to five percent, and five have six to eight percent. The highest percent for any county is 7.6 percent, the next lowest is 7.5 percent. The latter percent is that for the county under consideration. Of the eleven Minnesota cities having a population over 10,000, six have less than two percent of illiteracy, and three have from two to five percent. The percents for the remaining two are 7.7 and 16.1. The latter of these is for City "A" in this report.

These two considerations, that the mental test ratings do not do justice to the ability of the rural child, and that these particular schools have an unusually large percentage of cases of low mentality, are of importance. The third also bears considerable weight. Results from rural schools would be expected to be somewhat lower than those for city schools. That the difference is so great is indeed unexpected. As has been mentioned, 9.5 percent of the rural school cases fall below an I. Q. of 70. The percent of cases falling below 65 is 4.5; below 60, 1.5; below 50, .4. These percentages seem significant in the light of other surveys. It has been maintained that investigators who have depended upon social agencies, public officials, etc., for their information in regard to cases have discovered only the lower grade cases. That the percent of cases below an I. Q. of 50, i.e., imbecile or idiot-grade, found in these rural schools so nearly agrees with the results of surveys conducted in the above-mentioned manner seems to substantiate this belief.

In one or two surveys where tests of various kinds have been used as a partial means of diagnosis, it is evident, that in case the test result has not agreed with the investigator's personal judgment of the case, the latter has had the greater weight. There can be no question, that in such cases, a large number of the higher-grade feeble-minded, who are so difficult to classify, even for an expert in this particular field, have been dropped as not feeble-minded. It has been stated that this would be true in the absence of any mental rating. It seems a safe guess to say that the majority of cases, especially children, above an I. Q. of 60 or 65 would not be classed as definitely feeble-minded unless characterized by physical stigmata or unfavorable temperamental traits, by anyone not especially trained and experienced in the field of mental diagnosis, particularly when these cases are seen in surroundings where standards of achievement are not so high. It is alto-

gether possible that these same cases, when found in a rural community, and when judged by the social criterion, would not necessarily be classified as feeble-minded, as this term is legally applied in Minnesota.²² It is certain that a much smaller percent of these cases will need institutional care or state guardianship than of the cases of the same mentality in city schools. The 1.5 percent rating below an I. Q. of 60 in the rural schools seems to represent the group which would be classified as unquestionably feeble-minded had these schools been surveyed in the manner which has been indicated above. The 4.5 percent rating below 65 may be considered to comprise those cases most likely to become socially mal-adjusted and incompetent to maintain a livelihood independently. The 10.7 percent rating just above these are very likely the group which would not ordinarily be classed as feeble-minded because of their apparent normality when children, and because of the fact that in most instances, the simplicity of their environment and the relative ease of earning some sort of a living do not make greater demands on the adults of this group than their limited intelligence can meet.

The social investigations have been made for about one-fourth of the cases rating below an I. Q. of 75. The personal and family history obtained and the prevailing home conditions have confirmed the mental ratings in practically every case. In no case was there entire absence of significant features in the additional information thus received. It seems significant that in most of the schools surveyed, the teachers, a number of whom had taught in other parts of the county and state, did not recognize that they were dealing with an abnormal situation, or that they had cases of mental deficiency in their schools, although questioning usually brought out the fact that these children were considered 'peculiar,' or 'very slow,' or 'hard to understand.'

THREE SURVEYS COMPARED

The results for the rural schools of "X" County and those for City "A" of this county have been given separately. The combined results for these schools yield an average of 8.8 percent of cases below an I. Q. of 75. As this is drawing the line higher than was the case in the surveys of the counties in California and Indiana, the percent of cases below an I. Q.

²² "The term 'feeble-minded persons' in this act means any person, minor or adult, other than an insane person, who is so mentally defective as to be incapable of managing himself and his affairs, and to require supervision, control and care for his own or the public welfare."

of 70 was found. City "A" had 2.6 percent of its cases below an I. Q. of 70, the rural schools 9.5 percent, with an average of 6.1 percent for all the schools combined. These latter percents are comparable with those of Drs. Terman and Book. Dr. Terman classified cases as definitely feeble-minded in case they were below an I. Q. of 70. Dr. Book has not stated where he drew the line in classifying his cases, but since the Stanford Revision was used for individual examinations, it is presumable that it was not drawn above 70.

TABLE V

	City	Rural	Average
California	3.5	4.9	4.24
Indiana	2.	6.4	
Minnesota	2.6	9.5	6.1

The results for these three surveys are given in Table V. Unfortunately the average percent for the rural and city schools is not given for the Indiana Survey. From the two percents given, it can be estimated at about four percent. That the results for these three surveys should agree with each other much more closely than they do with the results of any of the other surveys reviewed above would be expected, as there are several points of similarity in the methods employed. The chief of these points of similarity are: (1) an accurate objective standard was used to diagnose mental deficiency; (2) a sufficiently large percent of individual examinations was made to warrant the assumption that all cases of mental deficiency had been found; (3) the surveys were under the direction of specialists in the field of psychology and mental diagnosis; and (4) the examinations were made by individuals trained in psychology and in the giving of mental tests. The chief difference in the methods was in the selection of cases for individual examination, group-tests not being used in the California Survey, nor in grades 1 and 2 in the Indiana Survey. Because of the above-mentioned points of uniformity in procedure, it seems safe to assume that the variation in the percents of feeble-mindedness found in these studies is indicative of the extent to which the localities differ in their respective proportions of mental deficiency.

SUMMARY

1. There is much lack of agreement in the results of the various surveys made to determine the percent of feeble-minded in the general or school population. The actual percentages quoted in the surveys which have been reviewed vary from .16 to 6.4.

28- - - - -

2. Most of this lack of agreement is due to the differences in the methods of conducting the surveys and to the variation in the methods and criteria used to diagnose feeble-mindedness.

3. Surveys of school systems in which the teachers have selected the cases suspected of feeble-mindedness give a percent of 2.2 of the total school enrollment which rates as feeble-minded. Teachers' errors are in two directions, selecting children who do not classify as feeble-minded and overlooking others who are feeble-minded. The chief cause of these errors is the failure of the teacher to take into consideration the age-grade status of the child.

4. The most complete type of survey is that in which carefully standardized group tests are used as a means of selecting all cases which are feeble-minded, these cases being further examined with individual mental tests before a final diagnosis is made. Surveys of this type in three cities of Minnesota yield an average of 4 percent of the number given group tests rating below an I. Q. of 75 in an individual mental test. With this method there is much less error in the selection of cases for individual examination than in the case of the teachers' selections.

5. The group test survey also makes possible a comparative study of the different grades and buildings in a system and reveals large discrepancies between the average levels of ability. For instance, the average I. Q. of the poorest grade in one building was found to be two points higher than the average I. Q. of the best grade in another building in the same system. A study of the distribution of the most superior and inferior in ability among the different buildings is also possible and likewise brings out striking contrasts. In one system of four buildings, one building had 60 percent of the most superior children of the entire system.

6. Mental surveys of thirty-five rural schools yield an extremely high percentage of cases rating below an I. Q. of 75. These schools had on the average 15 percent of the number present for the group test who rated below an I. Q. of 75 in an individual mental test. This percent was reduced to 9.5 when the upper limit for arbitrary classification of cases as feeble-minded was placed at an I. Q. of 70 instead of 75.

7. Combined results for the rural schools and City "A" of this county yield an average percent of 8.8 rating below an I. Q. of 75, and an average percent of 6.1 rating below an I. Q. of 70. These findings approximate those of other surveys employing similar methods.

ON FINDING EQUIVALENT SCORES IN TESTS OF INTELLIGENCE

By EDWARD L. THORNDIKE, Institute of Educational Research, Teachers
College, Columbia University

With the development of group tests and tests for use with higher levels of intelligence, it is becoming more and more necessary to transmute a score obtained with one test into the score that is equivalent to it in some other test, or into a mental age, which is, as now used, itself really an arbitrary score. We have already three tables of the mental ages corresponding to Alpha scores, one set computed by Yerkes and Yoakum ('20, pp. 133 f.); another, computed by Kohs and Proctor (Proctor, '21 pp. 65-68); a third, by DeVoss ('22, p. 24). Alpha and Beta scores have been equated over a part of their ranges. (Psychological Examining in the United States Army. Part III, Chapter II).

The procedure has been to obtain empirically the average score obtained in one test by those of a certain score in the other, and to use the series of such averages, smoothed to eliminate chance irregularities, as the table of equivalents. That is, the regression line is used. This is the natural thing to do and is probably, often, the most practicable thing to do, but it sometimes subjects the results to an important error. Theoretically it is the right thing to do only in special cases; and when it is done the resulting table of equivalents has a special meaning, which is not, I think, the meaning usually attached to such a table by most of those who will use it. It seems desirable therefore to consider the general theory of the determination of equivalent scores.

Case I. Consider first the case where the two tests measure the same thing, as when an inch scale and a centimeter scale are used to measure length, or when two adequate but different tests measure color blindness. If two tests measure the same thing and if each is perfectly reliable, correlating 1.00 with a repetition of itself, the empirical equating of values described above will be satisfactory, of course, *but there will be no regression*. The two tests will correlate perfectly. All the scores in Test II made by those of identical scores in Test I, will be identical.

In practice, of course, there will be a slight unreliability in each, due to the errors that affect all measures, but this will be, say, one part in a thousand, leaving a correlation between the two tests of say .999+. If a random sampling of lengths from 1 inch to 1 yard were measured by two of our most approved methods at the Bureau of Standards, the correlation would, I fancy, be above .99999 and the regression would be negligible in proportion to the magnitudes measured.

Case II. As Case II consider two tests which are known to measure the same thing,¹ but with a considerable unreliability. Such would be single estimates by the unaided eye of each of such a series of lengths first by one physicist in tenths of an inch and second by another in millimeters, or in a less degree the same measurement by two approved methods, but using only a single determination of each value, and letting these values be subject to errors of a careless copyist. Each of the two measures would then correlate imperfectly with itself, and one would correlate imperfectly with the other. The empirical comparison of estimates of length by the unaided eyes of the two physicists would give a substantial regression. 1 inch would "equal" more than 2.54 cm.; 36 inches would "equal" less than the (36) (2.54) cm. which it really equals. In such a case the regression line is obviously not a safe basis for equating the two scores. If the other regression is used 2.54 cm. would be found to equal more than an inch, while 91.44 cm. would equal less than a yard. In such an experiment with lines from 70 mm. to 90 mm. long, the estimated lengths were such as to make 2.5 inches equal to 68.7 mm. (5.2 mm. too long) and 3.8 inches equal to 91.7 mm. (4.8 mm. too short)!

What the regression line measures in Case II is the "attenuation," the combined unreliabilities of the two tests, the errors. If we had only such data to equate inches and centimeters by, we ought not to take the empirical equivalents as they stood but to correct them by some allowance for their "attenuation" or errors.

Suppose, for example, we know that Stanford Mental Age and Army Alpha measure the same thing. We plot the scores of one against the other for some thousands of cases using a single trial with each. If the correlation is not perfect,—

¹ "To measure the same thing" is somewhat ambiguous. The fact that the two measures do not correlate perfectly may be said to prove that they do not measure the same thing, length, but different things, to wit, length plus certain different causes of errors. This may be admitted. What I mean by the phrase here is, I think, clear.

if there is regression from perfect toward zero correlation, we must infer that this is due to our having imperfect measures in Stanford Mental Age or in Army Alpha or in both. A table of values of Alpha scores in terms of Stanford Mental Age got by taking the averages of arrays in Mental Age for each Alpha score—the regression line—smoothed or unsmoothed, is not a table of *equivalents of the scores*; it is, simply a table of *prophecies* of how a person who scores a certain amount in one trial with Alpha will probably be rated in Mental Age if he is given one trial with the Stanford test. It is as if we measured a man's stature in inches by having one person guess that it is 180 centimeters and looking up the average guess in inches made by another man when judging a 180 cm. height.

We can obtain equivalents only after first correcting for the attenuation. This can be done by applying Kelley's formula

$$(\sigma_{\infty} = \sigma_{\text{obt.}} \sqrt{\text{rob. 1. with obt. 2.}} \quad [21, \text{p. 379}])$$

to the Alpha series and the Stanford series. Then, knowing that the correlation really was perfect, we should fit the two corrected surfaces of frequency one on the other and use the correspondence of Alpha and Stanford that makes the best fit. If the attenuations in the two series were equal we could use the values which made the best fit of the two uncorrected surfaces of frequency, one on the other. If the two attenuations are equal, or if we know how the attenuation is divided between the two series, it should be possible to estimate the true equivalence from the two regression lines in this case, (the regression of Alpha on Stanford and of Stanford on Alpha), if we know that some one point in the relation represents the true equivalence and know which point it is. It is often reasonable to assume that the point where the two regression lines cross does approximately represent the true equivalence. It is at any rate the probably best representative of it that the data afford. A convenient way to do this is, if the two attenuations are equal, to plot the data on a scale such that 1 S. D. of A is the same length as 1 S. D. of B and then draw the bisector of the angle between the two regression lines.

I am not competent to derive a formula for the general case but I hope this article may stimulate someone to do so, at least for the case where the correlation is linear. The division of the attenuation between the two variables would be known either through the correlations (r_{a_1} with a_2 and r_{b_1} with b_2), or through the errors of the determinations, (e.g.,

the average σ true—obtained for a and the average σ true—obtained for b).

So the problems stand as follows:

A.

Given correlation plots (linear) of $r_{a_1b_1}$, $r_{a_1a_2}$, and $r_{b_1b_2}$
(Or in place of the two last given the average mean sq. error of the a 's and the average mean sq. error of the b 's)
Assuming that $r_{a \infty b \infty} = 1.00$
To draw the regression line for $r_{a \infty b \infty} = 100$
from the regression lines for $r_{a_1b_1}$

B.

The same, but for a correlation plot of any form for $r_{a_1b_1}$.

Case III. Consider now Case III, where we know that the two measures each have perfect reliability,—perfect self correlation. If then, in the empirical pairing of values, we find a lack of perfect correlation, we know that they *do not measure the same thing*. The regression from perfect correlation is then a measure of the unlikeness of the two things measured by the two tests. We ought not then to speak of one as having an equivalent in the other. The case is then logically like estimating a man's weight from his height. We can make a better estimate of a man's weight from knowing his height than from not knowing anything about him, but 6 feet does not equal *any* number of pounds. It is both logically and in fact like estimating the total wealth of an individual (a) by the taxes he pays and (b) by the amount he can borrow. Both a and b would correlate with the fact to be measured, but imperfectly, since each would measure different aspects of it. The wealth he pays taxes on is not, in every dollar, identical with the wealth he borrows on. It is almost certain that verbal tests like Alpha measure something which is in part not measured by a non-verbal test like Beta. Consequently, if a million persons were each measured by average score in a dozen Beta tests and by average score in a dozen Alpha tests so as to get almost perfect reliability for each, the two scores would still not correlate perfectly. A certain amount of Alpha ability would not then *equal* any amount of Beta ability. We could use one to *prophecy* the other, as in Case II.

There is a difference between the prophecy made in Case II and that made in Case III. We improve the prophecy in Case II toward perfect equivalence by making each test more reliable, for example, by repeating it several times, and we can infer what the true equivalents will be by the changes that occur with each known improvement in reliability. If

disturbances from practice effects, ennui, etc., are not too great, we may bring any two tests that measure the same thing, (for example, ability to spell, or general intelligence, or accuracy of vision) to a real equivalence. We can improve the prophecy in Case III only by altering the tests to make them measure more nearly the same thing. Nothing that we can do will bring the mean square error of estimate of an adult man's weight from his stature down to one pound. We must use more than stature to do that.

Case IV. Consider finally Case IV, where we have two tests which do not correlate perfectly, partly because neither test would correlate perfectly with a second trial of itself and partly because, even if we had hundreds of trials of each, the two averages would not measure quite the same thing. Before speaking of equivalent scores we must alter the tests until they measure the same thing and then correct the empirical equivalents so as to free them from the influence of "attenuation."

At the present time the commonest problems will be to determine the Stanford Mental Ages or the Army Alpha scores equivalent to scores in various group tests. These problems present Case IV with the added complexity that we do not know how much of the reduction from perfect correlation is due to the tests being measures of different traits, and how much to their respective "attenuations." There is, however, no easy way out of the difficulty. True tables of equivalents cannot be computed until we know the unreliability of each test, and the correlation between the two tests when freed from their unreliability (by increased length, repeated measurements, or otherwise). Unless the correlation is then perfect we cannot compute genuine equivalents of units of measure, but only prophecies of scores obtained by persons.

The facts presented here hold, of course, when the two tests to be equated are put in relation by the relation of each to a third variable, such as chronological age. In fact, the error due to taking the empirical averages of arrays (i.e. the regression) is increased in so far as the chronological ages are subject to errors or are grouped coarsely, as by years or half years.

REFERENCES

- DE VOSS, J. C., '22. *Manual of Instruction for Use with the Army Alpha Intelligence Tests*. Third edition.
 KELLEY, T. C., '21. The Reliability of Test Scores.
 PROCTOR, W. M., '21. The Use of Psychological Tests in the Educational and Vocational Guidance of High School Pupils. *Journal of Educational Research Monographs*, No. 1.
 YERKES, R. M., and YOAKUM, '20. *Army Mental Tests*.

SOME FACTORS OF AESTHETIC JUDGMENT

INTRODUCTION

By Mrs. ANNE ROSS REYMERT

The following experiments were undertaken between January and June, 1916, with a view of getting some light upon the various factors entering into aesthetic judgment.

After various preliminary tests upon miscellaneous observers, under different conditions, had been undertaken, the material, method and procedure for the main tests were determined.

In order to get at least a rough outlook over some of the main factors in aesthetic judgment, it was attempted to isolate, or to get these factors to *stand out*, one by one, in each of a series of different aesthetic preference-arrangements, thereby giving a means of comparison of the different rôle the specific factor may play, and its interplay with other factors in simple and more complex aesthetic judgments.

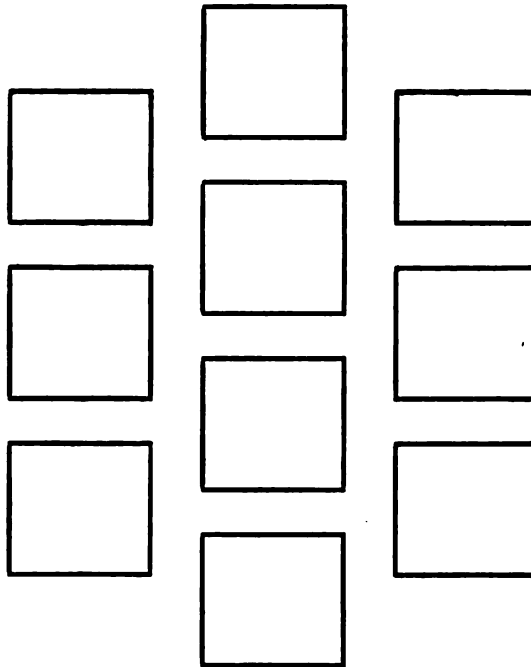
Twenty-two different series of aesthetic values were selected, each with the view of being:

1. a representative group series with reference to one specific aesthetic value, and
2. a comparable group in individual pictures (or card) value.

The observers consisted of two groups, one of 12 girls and the other of 12 boys, and were chosen at random among university students, mostly sophomores, juniors, and seniors, with two graduate students. All observers were without experimental training. The experimental conditions were kept uniform, and no tests were undertaken unless a subject reported himself in good condition.

The observers were experimented upon individually. The pictures were placed in a group arrangement as shown on inserted card. They were distributed within this plan so that no picture's relative position to any other picture might influence the observer's judgment, i.e., the scheme was to thoroughly mix up the pictures within the plan of arrangement.

This same arrangement, which afforded each picture an equal stimulus value, then was carefully maintained for all observers throughout all experiments. From Series S to Series EE, which involve the more complex stimuli, reasons



for the chief points upon which choice was based were requested.

Each observer's attention was carefully and explicitly directed to the one particular value to be judged in each series, and the instructions were kept uniform.

The observers were kept from communication with each other until tests were over.

The stimulus-cards or pictures were labelled with numbers on the reverse side where possible; otherwise they were lettered in a top-corner with a view of being unnoticeable and without meaning to the observers. The proportion tests, were labelled according to increase in width. To avoid influence of black and white effects in the proportion tests white outlines on black background were used.

Method

From the sum of the positions assigned to each individual picture by each group of subjects the average position or the average ranking of the picture was determined. And from

this the various correlations were calculated, by the method of least squares or the method of rank-differences. The correlation is thus a measure of the relative agreement of the judges with the average judgment of the group. The various additional tables are self-explanatory.

The lack of absolute standards in any preference judgments has always been a serious handicap. The above gives perhaps the best relative standard obtainable and has already been used advantageously by various investigators in the psychology of judgment. The results thus far derived seem to indicate that the method is not only adequate for all purposes of comparison, but gives great hopes for the future in various fields of psychology hitherto unattacked by experimental methods on account of the above-mentioned lack of an absolute standard.

To get ten homogeneous, comparable components that could make up the several series for the various aesthetic values to be judged, involved the prime difficulties in the selection of the materials for experimentation. The devising, constructing and organizing of the apparatus possible for use as a means of experimenting upon the *one particular* aesthetic element in isolation offered no less difficulties and required infinite exactness and most painstaking care in forestalling all possible sources of error and digression of the subject from the value exposed for his attention. To check these up a number of preliminary tests were made upon miscellaneous subjects.

Various art schools were drawn upon for available and possible material, suited to the testing out of the various aesthetic values, for which no apparatus was constructed. More had to be discarded than could be selected on account of the above mentioned prime difficulty.

THE VARIOUS EXPERIMENTS

The first three experiments consisted of the charcoal drawings group, Series A. B. and C. These particular drawings were mounted on a black screen in a similar general arrangement as mentioned above for all material used throughout these experiments.

The room for experimentation in this group was absolutely black, having no color effects in it whatever, and the only light that entered was the daylight coming from several small high windows. The screen was placed so that the light fell directly upon it at an angle of about 45 degrees. The observers were requested to seat themselves comfortably before the screen and contemplatively study these three series in

order, focusing their attention in Series A. upon the technique, i.e., excellency of drawing, in Series B. upon preference of pattern design, and in Series C. upon the most pleasing suggestiveness of a concrete object,—thus basing and noting their judgments upon one specific factor in turn for each series.

The following table may serve as a rough map over the substance of the problem, material, etc.:

Tests: 22 Series, each of 10 comparable aesthetic values respectively; judged by same 12 subjects.

Respective single factors as
bases of judgment—i. e.,
factors tested

Technique

Simple preference

Suggestiveness of concrete object

Kinds of Materials

Drawings of similar degrees of merit of
same study. (2 simple pieces of pot-
tery) drawn from slightly different
angles.

Drawings of pattern designs—each dif-
ferent.

Drawings of swirls—each different.

Photostats—white outlines on black
background mounted on white cards—
all of equal size.

Seven representative types: of non-geo-
metric figures, consisting of varying re-
lations between two curved lines and two
straight lines.

Each series having height of its figures
equal and equally graded increments in
width from 1 to 10.

Series V and VI have the same figure but
VI has a narrower proportion in width
for card 1 and a wider proportion for
card 10, and consequently larger in-
crements in width serially than Series
V.

No.	Series	Groups
1.	A.	<div style="display: flex; align-items: center;"> <div style="margin-right: 5px;">{</div> <div style="display: flex; flex-direction: column; gap: 5px;"> <div>Charcoal</div> <div>Drawings</div> </div> <div style="margin-left: 5px;">}</div> </div>
2.	B.	
3.	C.	

Proportion tests:

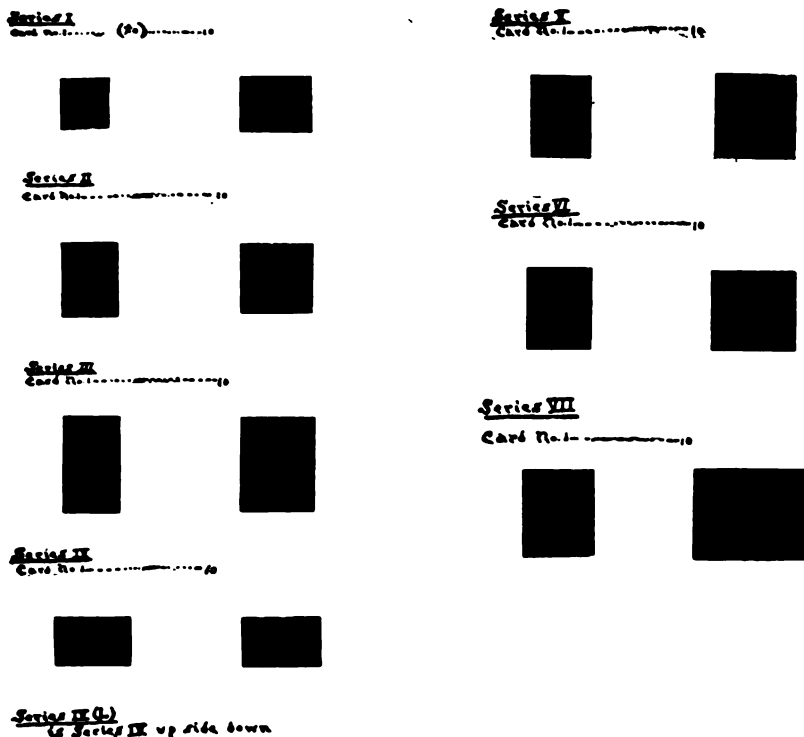
4.	I.		Width of entire figure increased.	Simple preference
5.	II.		Width of entire figure increased.	Simple preference
6.	III.		Width increased only at middle of figure.	Simple preference
7.	IV.		Width increased excepting at top.	Simple preference
8.	IV(L).		Width increased excepting at bottom.	Simple preference
9.	V.		Width increased only at top of figure.	Simple preference
10.	VI.		Width increased only at top of figure.	Simple preference
11.	VII.		Width of entire figure increased	Simple preference
12.	A & B		10 Arbor and Bird Day Magazine covers	Simple preference
13.	A & B (Color)	Same magazine cover illustrations	(Colored)	Preference based on color only
14.	A & B (Design)		(Colored)	Preference based on design only

For proportion figures
See p. 7 (a)

Proportion Tests

15.	D.	Ceramic art designs mounted on cards of same size.	<i>Non-geometric</i> All have same general proportions and same five colors in different arrangements.	Simple preference
16.	I.	Ceramic art designs mounted on cards of same size.	<i>Geometric</i> Designs with outstanding feature of complications of the lines—same five colors as in D.	Degree of intricacy of design
17.	S.	Illustrated Spring advertisements of similar size and mounted on cards of same size.	Non-colored. Word-matter removed.	Degree of suggestiveness of the idea of Spring
18.	AA.	Dutch Costumes.	Nine scenes representing Dutch costumes—colored.	Picturesque effect.
19.	BB.		Colored reproductions of paintings.	Preference with regard to desire as permanent possession in subject's room
20.	CC.	Post card reproductions.	Ten portraits of women. (Photogravures)	Most effective portrayal of feminine characteristic.
21.	DD.		Ten portraits of men. (Photogravures)	Most effective portrayal of masculine characteristic
22.	EE.		Ten religious pictures. (Photogravures)	Degree of impressiveness

PROPORTION FIGURES REDUCED ONE-HALF



EXPERIMENT I

Series A. Drawings of similar degrees of merit of the same study, two simple pieces of pottery, drawn from slightly different angles. The judgment was based upon the technique, or excellency of drawing.

Correlating in *Series A.* each individual girl judgment with the average of the girl group ranking the following correlations were found:

(Small letters refer to girls and capital letters to boys)

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	eh	gs
0.30	0.14	0.43	0.75	0.65	0.55	0.48	0.37	0.71	0.68	0.38	0.49

Average of above correlations equals 0.49.

With an average deviation of plus or minus 0.14.

The same results for the boys:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.59	0.25	0.90	0.45	0.70	0.88	0.62	-0.30	0.90	0.24	-0.13	0.40

Average of above correlations equals 0.46. —

With an average deviation of plus or minus 0.31.

Judging Series A. then, on the basis of technique, the girls agree slightly more within their group than do the boys. Also their variation around the group judgment is only half as great as among the boys.

The average correlation between the sexes as a group in this series, (0.49), shows the general agreement of position assigned by the two groups. They differ markedly only in regard to the place assigned to one item, which is ranked third by the girls and tenth by the boys. The group correlation for the remaining nine items is 0.71.

The table below indicates the factors determining the average ranking of each individual picture for the girls.

Sum of Positions Assigned	Average Ranking	Average Position	Average Deviation of Pictures
88	9	7.3	plus 2.3
71	5	5.9	or — 1.5
22	1	1.8	1.2
62	4	5.2	2.4
75	8	6.3	2.4
48	2	4.0	2.0
73	7	6.1	2.4
72	6	6.0	2.2
61	3	5.1	2.1
88	10	7.3	1.6

Pictures ranked number 1, 10 and 5 clearly show the smallest average deviations of position. Picture ranked number 1, owes its position through being ranked first by 9 of the 12 observers. Picture number 10, however, owes its position to a conflicting set of judgments, i.e., to spread votes. Similarly, picture ranked number 5 also owes its position to spread votes. From this the conclusion follows that clearly the greatest degree of agreement falls for the girls on the picture most preferred by both girls and boys.

Similarly, the following table for the boys:

Sum of Positions Assigned	Average Ranking	Average Position	Average Deviation of Pictures
72	6	5.8	plus 2.2
78	7	6.5	or — 1.7
23	1	1.9	0.9
62	4	5.2	1.8
81	9	6.8	2.5
52	2	4.3	2.4
62	3	5.2	2.7
66	5	5.5	2.0
87	10	7.3	2.5
77	8	6.6	1.9

Thus, for the boys it is obvious that the greatest degree of agreement likewise falls on the picture most preferred.

EXPERIMENT II

Series B. Charcoal drawings of pattern design. Here the judgment was based upon simple preference.

Labels of pictures	Average ranking by girls as a group	Average ranking by boys as a group
1	10	10
2	3	1
3	9	8
4	5	3
5	2	2
6	8	7
7	1	5
8	4	6
9	6	4
10	7	9

The individual correlations of the girls with their average ranking as a group are as follows:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
0.77	0.43	0.76	0.71	0.78	0.78	0.84	0.59	0.82	0.71	0.76	0.77

The individual correlations of the boys with their average ranking as a group are as follows:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.89	0.55	0.52	0.76	0.89	0.13	0.98	0.85	0.44	-0.19	0.08	0.21

Girls:

Average correlation of
individuals with group
order is 0.73
Average Deviation equals ± 0.08

Boys:

Average correlation of
individuals with group
order is 0.51
Average Deviation equals ± 0.31

Correlation between girls and boys as one group equals 0.77.
In judging these charcoal drawings of pattern design on

the basis of simple preference the following results then are obtained.

As a group the girls agree very highly and show little individual variation around the group judgment.

The boys correlate only moderately high in this experiment, —considerably less than the girls. However, their span of variability is four times greater.

EXPERIMENT III

Series C. Charcoal drawings of swirls. Judgment based on: most pleasing suggestiveness of a concrete object. Here then, a more complex factor was involved.

Labels of pictures	Average ranking by girls as a group	Average ranking by boys as a group
1	5	3
2	4	4
3	3	5
4	10	9
5	9	7
6	2	2
7	8	6
8	1	1
9	6	10
10	7	8

The individual correlations of the girls with their ranking as a group are as follows:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
0.75	0.72	0.70	0.08	0.75	0.49	0.88	0.59	0.81	-0.31	0.77	0.88

The individual correlations of the boys with their ranking as a group are as follows:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.71	0.72	0.72	0.82	0.62	0.41	0.62	0.38	0.41	0.62	-0.36	0.64

GIRLS:

Average correlation of
individuals with group
order equals

Average Deviation ± 0.25

Boys:

Average correlation of
individuals with group
order equals

Average Deviation ± 0.21

Correlation coefficient between girls and boys as one group equals 0.79.

Here again as a group, the girls agree more than do the boys; however, here they show, contrary to the former experiment, a higher degree of individual variation.

The attitude of the observers to Series C. was distinctly contemplative and the response in judging it distinctly more weighed and considered than in the previous experiment. From the remarks of the observers about their mental setting it was

apparent that the imaginative reaction played an insignificant part compared with the aesthetic reaction. In the imaginative reaction *no* judging factor was involved, hence the response much more immediate. However, with the entrance of the aesthetic reaction a critical, selective element entered, which was readily apparent from the comments of the observers upon the material. The dual process involved first, from the conditions of the test, that the aesthetic reaction be preceded by imaginative reaction and checked up by recognition. The characteristic then, of the aesthetic reaction seems to be distinctly contemplativeness.

Comparing the qualities of the aesthetic reaction in the three experiments, we may designate the first, Series A., as both distinctly and distinctively *objective* aesthetic reaction, and Series B. and C. purely *affective*; B., however, being simpler, and having less of the critical, is therefore, to a lesser degree contemplative.

From the comparison the average of group correlations for all these three experiments:

	Girls	Boys	Both Sexes as One Group
Series A	0.49	0.46	0.49 (or 0.71 for
Series B	0.73	0.51	0.77 nine pictures)
Series C	0.59	0.53	0.79

it seems, in contrasting the affective (B. and C.) with the objective (A.), as if there were greater agreement in more affective aesthetic judgments, both within the sex groups and between the sexes as one group.

In the above experiments there seems to be an increase of agreement between the sexes, with increase of the complexity.

Taking the average correlation of the group correlations, for girls, boys, and the two sexes together, from the above three experiments:

Girls	0.60	Boys	0.50	Both Sexes as One Group	0.68
-------	------	------	------	-------------------------	------

it appears, that in aesthetical judgments based on technique, simple preference, and pleasing suggestiveness of a concrete object, the girls will approach more closely the mixed-group standard than will the boys, presumably because their preferences are less widely spread, i.e., more closely concentrated.

EXPERIMENTS IV TO XI

Proportion Tests (for materials see outline above).

Judgment based on simple preference.

SERIES I TO VII

Average rankings in all proportion tests of girls and boys as separate groups:

REYMERT

CHART X

Labels of Cards	Series I		Series II		Series III		Series IV		Series IV(L)		Series V		Series VI		Series VII	
	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B
1.....	10	7	8	10	8	10	10	10	8	10	9	7	9	10	10	10
2.....	8	9	5	9	6	8	8	7	2	7	7	6	8	9	9	8
3.....	5	3	6	7	1	4	6	6	4	5	5	2	7	8	8	7
4.....	3	6	3	5	2	2	5	3	3	6	3	5	1	7	7	6
5.....	1	1	2	2	4	3	4	2	7	4	4	1	4	4	4	1
6.....	2	2	1	1	5	1	1	1	1	1	1	4	3	2	2	1
7.....	4	4	4	3	3	5	3	4	5	6	8	10	2	6	1	2
8.....	6	5	7	6	7	7	2	5	9	8	6	7	7	2	3	3
9.....	7	8	9	4	9	9	7	8	10	9	10	9	10	8	5	4
10.....	9	10	10	8	10	9	9	9	9	8	9	9	8	8	6	4

CHART Y

Labels of Cards	Average Rankings	Average Positions	Average Deviations
1.....	10	8.8	Plus 1.3
2.....	8	7.3	or— 1.4
3.....	5	5.1	1.5
4.....	4	4.1	1.6
5.....	2	3.1	1.5
6.....	1	1.8	0.9
7.....	3	3.8	1.3
8.....	6	5.3	1.5
9.....	7	6.9	1.2
10.....	9	8.7	1.2

The above "Chart X" gives indications in reference to the *factors determining* the ranking of the cards, in all proportion tests.

Taking both sexes together, the greatest degree of agreement of actual votes falls both on the most preferred and on the least preferred card. Between these two extremes themselves, however, by far the greatest degree of agreement by actual vote falls on the most preferred card. The other cards owe their position chiefly to distribution of spread votes, but these—as may be seen from the average rank and also the average position of the cards in Chart Y—distribute themselves in regular increasing succession above and below the most preferred card. However, in glancing over the actual agreement of votes of the girls and of the boys separately, it is apparent that the boys agree more decidedly on both their most as well as on their least preferred cards, while the girls agree more emphatically on their least preferred card than they do on their most preferred card. The most preferred card of the girls owes its position more to spread votes than to actual agreement. Their least preferred card owes its position to their most emphatic agreement of choice. Considering the medium choices of each sex, the greatest actual choice among the boys falls on card labelled 3, and that of the girls on card labelled 8.

As *one group* of people then, the girls and boys together agree in the liked and disliked by actual votes, i.e., in the extremes.

As separate groups, the girls agree by actual vote only on their least preferred extreme; the boys by actual vote on both extremes, and on these choices only.

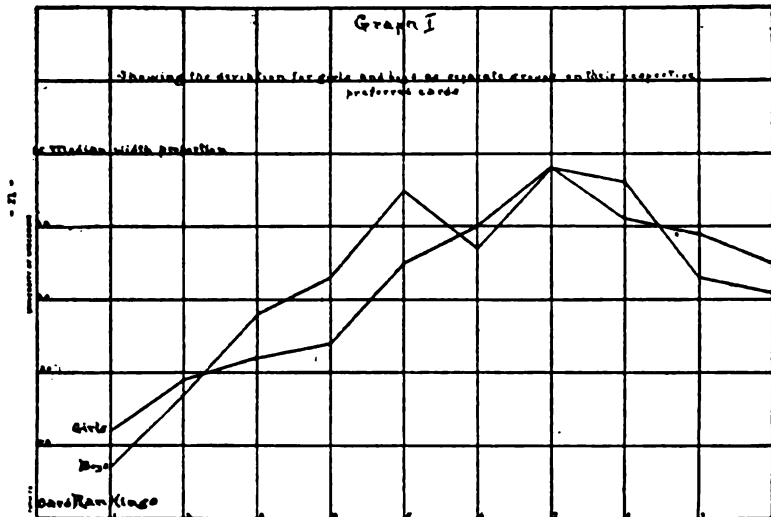
That is, together the sexes converge on likes but not on dislikes. Separately, the girls agree only on dislikes among themselves; the boys both in likes and dislikes.

CHART Z

Grouping the average ranking of the sexes together:

Card Labels	Sum Average Position	A. D.	Average Ranking	Average Ranking	Sum Average Position	A. D.
	Girls			Boys		
1.....	72	+17	9	10	70	+15
2.....	53	-2	7	8	64	+9
3.....	40	-15	5	6	42	-13
4.....	27	-28	3	4	38	-17
5.....	29	-26	4	2	22	-33
6.....	17	-38	1	1	12	-43
7.....	24	-31	2	3	33	-22
8.....	45	-10	6	5	40	-5
9.....	59	+4	8	7	53	-2
10.....	74	+19	10	9	66	+11

The most strikingly outstanding feature is the fact of great relative general average agreement between the girls and boys successively, reading from card labelled 1 downward. But starting from the most preferred card (labelled 6) there is a successive distribution toward extremes in both sexes, yet with a gradual successive divergence above and below toward the extremes.



As is readily seen, from Chart Z, the middle choices of the two sexes, i.e., choices five and six, fall on cards labelled three and eight, showing clearly that their medium judgments fall half-way between the most preferred and the least preferred card of each sex, i.e., cards labelled one and ten. Then it may be concluded that, as two separate sex-groups, the sexes agree on the most preferred and the medium preferred cards. Also, Chart Z shows that the distribution of the deviations follows successively the distributions of the choices; the most preferred card, (number six), having the greatest deviation; the medium preferred a lesser deviation; then cards number two and number nine having the least deviation, and cards one and ten again having a medium deviation. Also, Graph I clearly indicates the similar direction of the choices of the two sexes.

The fact that there is one *definite, single, most-preferred* proportion figure, the medium one, and the further fact, that again there is a midway point of agreement between this median figure and the two extremes of proportion differences suggest: since the cards were labelled successively from one to ten in accordance to increase in width, thus giving a regular rate of proportions between two curved and two straight lines, it may be inferred that there probably is a medium preferred non-geometric type of proportion between two curved and two straight lines. The boys were emphatic in their first choice by actual average agreement, and also placed this agreement into the wider half of the proportions. When it came to choosing between the narrowest and the widest extreme proportions as the least preferred one the girls and boys were equally emphatic in their choice, *but* so in diametrically opposite directions.

May there then be a medium type of curved and straight line combination figure calling out a uniform aesthetic appeal?

May there be a "golden" proportion of this type analogous to the golden proportion of straight lines, or to the line of Hogarth? The kind of aesthetic response it calls forth seems to be very similar. Are the basic elements in this response of physiological origin?

To give a general view over the correlative results in all proportion tests the following table is inserted. It substantiates in the main the above inferences.

REYMERT

INDIVIDUAL CORRELATIONS IN ALL PROPORTION TESTS

	mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
Series I (G.)	0.53	0.85	0.85	0.25	0.53	-0.27	0.61	0.04	0.26	-0.04	0.02	0.54
Series II (B.)	0.85	0.16	0.79	0.50	0.33	0.37	-0.02	0.28	0.50	0.14	0.38	-0.32
Series III (G.)	0.88	0.84	0.52	0.73	0.76	0.28	0.87	-0.32	0.32	0.95	0.33	0.41
Series III (B.)	0.71	0.43	0.49	0.62	0.90	-0.07	0.75	0.35	0.42	0.66	-0.36	0.48
Series IV (G.)	0.59	0.79	0.33	0.75	-0.05	0.84	0.77	-0.44	0.16	0.81	0.43	0.45
Series IV (B.)	0.62	0.42	0.76	0.81	0.53	0.81	0.58	0.49	0.87	0.13	0.13	0.47
Series V (G.)	0.38	0.84	0.10	0.77	0.30	0.68	-0.22	-0.27	0.37	0.96	0.65	0.98
Series V (B.)	0.54	0.37	0.90	0.48	0.60	0.71	0.84	0.26	0.20	0.82	0.53	0.60
Series VI (G.)	0.09	0.50	0.76	0.59	0.64	-0.45	0.65	0.47	0.84	0.65	0.78	0.25
Series VI (B.)	0.89	0.66	0.58	0.38	0.94	0.37	0.01	0.50	0.52	0.72	-0.37	0.18
Series VII (G.)	0.71	0.83	0.84	-0.03	0.49	0.10	0.05	-0.07	0.07	0.81	0.09	0.13
Series VII (B.)	0.82	0.66	0.62	0.10	0.41	-0.55	0.85	0.72	0.67	-0.19	0.19	0.36
Series VIII (G.)	-0.05	0.87	0.15	0.15	0.94	0.05	0.03	-0.05	0.05	0.92	0.12	-0.03
Series VIII (B.)	0.84	-0.36	0.82	0.78	0.28	0.81	0.84	0.61	0.94	0.07	0.37	0.55
Series IX (G.)	0.66	0.99	-0.20	0.98	-0.59	0.66	0.89	0.66	0.95	0.60	0.66	0.76
Series IX (B.)	0.70	0.81	0.28	0.90	0.89	0.77	0.14	0.78	0.88	0.70	-0.41	0.71
CC		CJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK

Series		Average Correlation	Average Deviation	Correlation be- tween Girls and Boys as one group
I.	(G).....	0.35	± 0.30	0.84
	(B).....	0.33	0.24	
	(G).....	0.55	0.29	
II.	(B).....	0.45	0.25	0.65
	(G).....	0.45	0.31	
III.	(B).....	0.55	0.19	0.73
	(G).....	0.46	0.35	
IV.	(B).....	0.57	0.17	0.88
	(G).....	0.48	0.26	
IV.(L)	(B).....	0.45	0.28	0.59
	(G).....	0.34	0.34	
V.	(B).....	0.39	0.34	0.33
	(G).....	0.26	0.32	
VI.	(B).....	0.55	0.30	0.70
	(G).....	0.59	0.32	
VII.	(B).....	0.60	0.29	0.94

The median correlation for all judgments of the girls is 0.44 ± 0.31 .

The median correlation for all judgments of the boys is 0.49 ± 0.26 .

The median correlation between boys and girls, as one group, is 0.71

Omitting, however, Series V (which involved the finest increments of width increase, and hence, on account of the fine discrimination required, the choices may have been determined largely by chance), we get a correlation between the sexes of 0.76.

The correlation in both of the sex groups is very low, somewhat lower for the girls than for the boys, and the deviation is almost as great as the correlation, the boys showing a somewhat lesser variability.

The span of the individual correlations for the girls is -0.59 to 0.99 , and for the boys is -0.55 to 0.94 , showing a similar range of great individual variability for both sexes.

The average deviations on the cards in Series V, as compared with those for all the cards in all proportion tests, are more than twice as great.

The average correlation between the sexes as one group in Series V is the lowest for all twenty-two experiments.

Labels of Cards	SERIES V			
	GIRLS		BOYS	
	Card preference or ranking	Average position of cards Average Deviation	Card preference or ranking	Average position of cards Average Deviation
1.....	9	6.6 Plus 3.8 or—	3	4.7 Plus 3.0 or—
2.....	7	5.8 3.0	6	5.0 2.0
3.....	5	4.9 2.4	2	4.5 1.4
4.....	3	4.9 1.1	5	4.8 0.9
5.....	4	4.9 1.3	1	3.9 2.3
6.....	1	3.9 1.8	4	4.8 1.6
7.....	2	4.8 1.8	8	6.6 1.7
8.....	8	6.4 2.8	10	7.6 1.9
9.....	6	4.8 2.4	7	6.3 2.8
10.....	10	7.4 3.0	9	6.9 3.0
Average correlation for the girls as a group = 0.34				
Average deviation = ± 0.34				
Average correlation for the boys as a group = 0.39				
Average deviation = ± 0.34				
Average correlation between the sexes as one group = 0.33				

In Series IV and IV (L) the same figure was used, in the first instance with the widened portion at the bottom, in the second, it was exposed upside down. The lesser correlation between the sex groups in the second case is perhaps accounted for by the fact that the increase of width of the figure may have been more in a direct line of vision in the first case, and that we put meaning into characters from the top downward, hence notice the increase in width more attentively in Series IV.

There is also a sex difference in these two series:

	Girls as a Group	Boys as a Group	Girls and Boys as a Group
Series IV—Average Correlation.....	0.46	0.57	0.88
Average Deviation.....	± 0.35	± 0.17	
Series IV(L) Average Correlation.....	0.48	0.45	0.59
Average Deviation....	± 0.26	± 0.28	

The more obvious increase in width in the first exposure seems to have affected the girls negatively, but the boys positively.

Roughly, there is in these tests the same general variability as in the previous charcoal group.

EXPERIMENT XII-XIV

Series	Judgment based on—
A & B	Simple preference
A & B—(Color)	Preference based on color.
A & B—(Design)	Preference based on design

For this group of experiments, illustrative, colored Arbor and Bird Day magazine covers were used and so arranged for exposure that especially the color values of each magazine cover had an equal chance of appeal.

From the average rankings in these three experiments, it is of prime interest to find the possible part color and design, respectively, have played in the first selection of the magazine covers, in which all factors had free play.

Labels of Pictures	AVERAGE RANKINGS					
	A & B Simple Preference		A & B Color		A & B Design	
	(G)	(B)	(G)	(B)	(G)	(B)
1.....	10	9	9	9	8	6
2.....	8	10	10	10	7	10
3.....	5	4	4	6	5	8
4.....	7	1	7	2	4	1
5.....	1	3	1	1	3	4
6.....	6	7	6	7	6	9
7.....	9	6	8	8	10	5
8.....	2	8	5	4	2	3
9.....	4	5	3	5	1	2
10.....	3	2	2	3	9	7

It is found that the correlation between *simple preference* and *color* is:

for the girls, as a group 0.89

for the boys, as a group 0.82

which goes to show that color in the general judgment has played a conspicuous part, much more so than design, as the correlation between *simple preference* and *design* is:

for the girls as a group 0.61 (if last picture omitted)
correlation = 0.77

for the boys as a group 0.45 (if last picture omitted)
correlation = 0.46

This substantiates the psychological primacy of *color to form* in the aesthetical development of the individual and the rôle it plays in aesthetic judgment.

Also the girls correlate higher, both in color as well as in design, with their general simple preference judgment, than do the boys.

The following table gives a clear oversight over the correlations of all three series independently:

The individual correlations of the girls with their ranking as a group:

Series:	mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ge
A. & B.	0.59	0.64	0.88	0.25	0.53	0.68	0.87	0.76	0.41	0.50	0.66	0.26
A. & B. (Color)...	0.55	0.66	0.27	0.28	0.40	0.68	0.48	0.81	0.82	0.81	0.84	0.73
A. & B. (Design).	0.75	0.90	0.43	0.58	0.92	0.49	0.81	0.64	0.67	0.13	0.79	0.31

The individual correlations of the boys with their ranking as a group:

Series:	CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
A. & B.	0.89	0.53	0.67	0.13	0.78	0.20	0.45	0.21	0.40	0.66	0.58	0.33
A. & B. (Color)...	0.62	0.54	0.07	0.05	0.54	0.04	0.27	0.47	0.50	0.59	0.60	-0.19
A. & B. (Design).	0.24	0.22	-0.01	0.79	-0.03	0.52	0.68	0.47	0.13	0.04	0.73	-0.35

Series		Average Correlations	Correlation between sexes as one group
A & B	G.....	0.59±0.16	0.43 (0.52 for 9 pictures)
Simple preference	B.....	0.49±0.20	
A & B	G.....	0.60±0.16	0.78
(Color)	B.....	0.34±0.25	
A & B	G.....	0.62±0.19	0.56
(Design)	B.....	0.29±0.29	

As is seen here, the girls correlate much higher within themselves than do the boys. The variations in the correlations of the girls and boys as separate groups, in these series, may suggest that girls have more of a general standard in such aesthetic judgments in which color and design are involved than the boys have. Obviously the sexes agree best on color preference.

EXPERIMENTS XV—XVI.

Series D: Judgments based on design—

Series I: Judgments based on intricacy of design—

The material here consisted of Spanish ceramic art designs, essentially decorative, with a similar basic motif and having the same five colors. The design in the first series were non-geometric.

AVERAGE RANKINGS

Series D.

Individual correlations of the girls with their average ranking as a group:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ga
0.77	0.50	0.77	0.82	0.71	0.75	0.76	0.67	0.70	0.78	0.22	0.77

Individual correlations of the boys with their average ranking as a group:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	DN	RG
0.83	0.59	0.59	0.78	0.88	0.41	0.40	0.82	0.38	0.14	0.75	0.47

Series I.

Individual correlations of the girls with their average ranking as a group:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ga
0.98	0.93	0.05	0.54	0.82	0.64	0.90	0.83	0.76	0.83	0.33	0.75

Individual correlations of the boys with their average ranking as a group:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.82	-0.07	0.26	0.98	0.87	0.95	0.92	0.89	0.33	0.48	0.09	0.64

Labels of Pictures	AVERAGE RANKINGS			
	Series D		Series I	
	Girls	Boys	Girls	Boys
1.....	2	4	10	9
2.....	10	10	7	10
3.....	5	7	5	8
4.....	6	8	9	7
5.....	1	1	4	6
6.....	3	2	6	5
7.....	9	5	8	4
8.....	4	3	3	2
9.....	7	9	1	3
10.....	8	6	2	1

It is found that between these aesthetic preference judgments of design as such, and judgments based on intricacy of design, the girls correlate negatively (-0.05) and the boys very low, (0.25), which fact may indicate that as soon as an intellectual element is carried into the aesthetic judgment the preference may change almost entirely,—away from a directly impressive, aesthetic response.

Series		Average correlations of girls and boys as separate groups	Average Deviation	Correlation be- tween girls and boys as one group
D	G.....	0.69	plus 0.11	0.77
	B.....	0.59	or—	
	G.....	0.70	0.19	
I			0.20	0.70
	B.....	0.60	0.32	

Correlation between girls as a group for both series: -0.05 .
Correlation between boys as a group for both series: 0.25 .

The greater variability in Series I as compared with Series D also makes us surmise that the entrance of the intellectual interferes with the full reaction of the aesthetic. The relatively equal correlations of the sex groups for both series indicate that there are great similar agreements within the individuals of the groups—for an entirely different aesthetic ranging—given the different conditioned judgment bases of Series D and Series I.

There is also great agreement between the sexes in this respect.

TO BE CONTINUED IN JUNE NUMBER

MINOR STUDIES IN THE PSYCHOLOGY OF ADVERTISING

From the Psychological Laboratory of Indiana University

V. THE DEVELOPMENT OF ART-FORMS IN MAGAZINE ADVERTISING

By HARRY DEXTER KITSON¹

It is evident to even the casual observer that art plays a much greater part in advertising than formerly. Some conception of this fact may be secured from figures reported in an earlier number of this series,² showing that the number of illustrations increased on the average from 26 in a hundred to 85 in a hundred during the quarter century from 1895 to 1919.

There should, however, be further differentiation in describing this development; for simultaneously with the increase in number of pictures has gone great advancement in kindred respects—photography, color process, etc. Furthermore, all apart from the mere matter of quantity has gone great advance in quality, until the advertisement of today is to that of thirty years ago as a pastel is to a chromo. This advance has come about by reason of the development in technique and the evolution of various art forms. It was to analyze these out and express their development quantitatively that this investigation was undertaken.

The data for the investigation were secured from the files of two representative weekly magazines—*Literary Digest* and *Harper's Weekly*. (The latter discontinued publication in 1916 and so *Collier's* was substituted after this date.) All the advertisements were examined (with the exception of those on the cover pages which had been removed in binding), for the years 1895, 1900, 1905, 1910, 1915, 1920. The illustrations were classified as follows: pen and ink, wash drawing, photograph, pencil, charcoal and combination (any one of these combined with another or with color).

¹ With the assistance of R. D. Bruce and G. M. Dykes.

² H. D. Kitson, Amount and Rate of Increase in the Use of Illustrations, *Jour. of App. Psychol.*, March, 1921, p. 12.

TABLE I
Showing the number and per cent of illustrations in the Literary Digest and Harper's Weekly from 1895 to 1920, Classified according to form of art used

	Pencil		Pen Ink		Wash		Photo		Combina- tion		Char- coal		Misc.		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	
1895															
Lit. D.....	0	0	1128	96	16	1.	37	3.	2	0	0	0	0	0	1182
H. W.....	16	1.5	878	85.4	59	5.7	65	6.3	5	.48	3	.29	1	.09	1027
Total.....	16	.7	2006	90.7	75	3.3	101	4.6	7	.3	3	.1	1	.04	2210
1900															
Lit. D.....	0	0	350	80.	80	17.	12	3.	3	0	0	0	0	0	445
H. W.....	17	1.6	580	57.1	176	17.3	151	14.8	84	8.2	2	.19	4	.39	1014
Total.....	17	1.2	930	63.8	256	17.2	163	11.1	87	5.9	2	.1	4	.2	1459
1905															
Lit. D.....	1	0	821	68.	177	15.	121	10.	79	7.0	1	0	0	0	1200
H. W.....	56	8.4	310	46.9	66	16.	130	19.7	80	12.1	3	.45	15	2.2	660
Total.....	57	3.0	1131	60.7	243	12.3	251	12.9	159	8.6	4	.2	15	.8	1860
1910															
Lit. D.....	6	.1	1353	41.3	947	29.	832	24.5	139	4.2	0	0	0	0	3277
H. W.....	42	6.4	325	49.7	115	17.6	78	11.9	90	13.6	1	.15	2	.3	653
Total.....	48	1.2	1678	42.7	1062	27.	910	23.1	229	5.7	1	.02	2	.05	3930
1915															
Lit. D.....	10	.2	1330	36.1	1136	27.4	642	15.2	1086	25.7	1	.0	4	0	4209
H. W.....	4	1.4	151	54.3	48	17.3	28	10.	45	16.2	2	.7	0	0	278
Total.....	14	.3	1481	33.0	1184	26.1	670	14.9	1131	25.2	3	.06	4	.09	4487
1920															
Lit. D.....	31	.6	1102	21.8	1402	27.8	1021	20.2	1483	29.3	3	0	5	0	5047
Colliers.....	77	3.6	448	21.	765	31.4	204	9.7	566	26.3	3	.14	13	.6	2076
Total.....	108	2.1	1550	21.5	2167	30.3	1225	17.1	2049	28.7	6	.08	18	.2	7123
Grand Total..	260	1.2	8776	41.6	4987	23.6	3321	11.	3662	17.2	19	.09	44	.2	21069

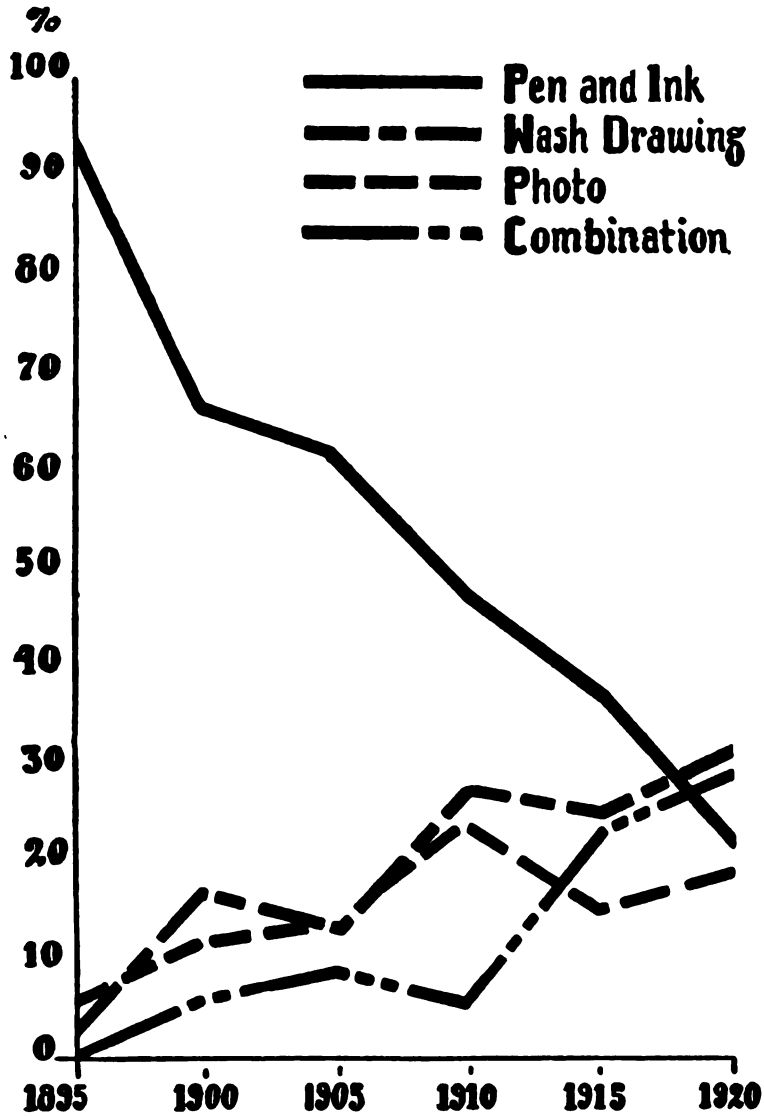


FIG. 1.—Showing per cent of illustrations in advertisements, 1895-1920, classified according to form of art employed

RESULTS AND CONCLUSIONS

1. The results show clearly that the early illustrations were done almost entirely through the medium of pen and ink, this constituting in the *Literary Digest* 96 percent, and in *Harper's Weekly* 85 percent, of the whole. (Photographs constitute the only other form in the early issues of the former.) As time went on, pen and ink work lost in relative popularity until in 1920 it constituted only 21 percent; though this amount still forms an important fraction of the whole—almost as large as that of any other art-form.

2. While pen and ink work has been waning, other forms have been coming to the front with varying degrees of rapidity:

a. The most notable advance has been made in wash drawings which have increased on the average from 3 percent in 1895 to 30 percent in 1920.

b. Photographs used alone have increased from an average of 4 percent in 1895 to 17 percent in 1920. There has really been a greater increase than is indicated by this, for many of the illustrations classed as combinations employed some photography.

It is of interest to note that the curves showing number of wash drawings and of photographs follow each other closely; this reflects the fact that the same technique of reproduction applies to both.

c. Pencil and charcoal have occupied so slight a place that the figures therefor are not presented graphically. Color, which in the present instance has been included within the combinations, involves factors so unique that its growth has been treated in a separate investigation (see below).

d. Combination illustrations, by reason of the development of technical skill in treating the forms just mentioned, and by reason of the increase in audacity on the part of all concerned with commercial art, have increased from .3 percent in 1895 to 28.7 percent in 1920.

These figures constitute a history, in quantitative terms, of the development of commercial art during the past quarter-century. In addition to the conditions they reveal concerning the advertising of the past, is the discovery that the several forms of art now enjoy about equal popularity.

It is of course understood that these figures may not serve with equal accuracy as descriptive of all advertising. Figures gathered from a study of other magazines might be slightly different, and those gathered from newspapers might differ more widely still. Still the facts here portrayed show the

TABLE II
Showing number and per cent of colored advertisements in four magazines, 1905-1920. (Files for certain years incomplete.)

	American			Colliers			Ladies Home Journal			Country Life			Totals		
	Full Page	Color Used	Per Cent	Full Page	Color Used	Per Cent	Full Page	Color Used	Per Cent	Full Page	Color Used	Per Cent	Full Page	Color Used	Per Cent
1905.....	0	0	0	x	x	x	0	0	0	901	10	1.1	901	10	1.1
1906.....	0	0	0	x	x	7.2	0	0	0	907	15	1.6	907	15	1.6
1907.....	0	0	0	563	44	7.2	0	0	0	815	14	1.6	1378	58	4.2
1908.....	0	0	0	513	34	5.6	35	1	2.7	660	11	1.6	1268	46	3.6
1909.....	0	0	0	439	11	2.4	54	6	10	846	25	2.8	1339	42	3.1
1910.....	0	0	0	644	18	2.7	26	5	16.1	774	7	.8	1444	30	2.0
1911.....	0	0	0	798	23	2.8	62	3	4.6	753	17	2.2	1613	43	2.6
1912.....	0	0	0	862	36	4.0	74	6	7.5	1076	28	2.5	2012	60	2.9
1913.....	189	49	27	980	17	1.7	80	3	3.6	901	28	3.0	2150	97	4.5
1914.....	134	61	45	838	28	3.2	64	4	5.8	983	19	1.9	2019	152	7.5
1915.....	110	44	44	956	78	7.5	41	24	37	787	10	1.2	1294	156	12
1916.....	204	61	30	1168	90	7.2	106	54	8.8	826	36	3.9	2304	241	10.4
1917.....	299	60	20	1224	64	4.9	177	138	43.8	768	36	4.4	2468	298	12
1918.....	414	72	17	915	50	5.1	192	208	52	558	34	5.7	2079	364	17.5
1919.....	715	122	16	707	42	5.6	317	391	53.2	671	39	5.4	2410	594	24.6
1920.....	959	221	24	652	93	12.4	956	481	50	382	97	20.2	2949	892	20.2

trends significantly, and offer a factual instead of a speculative basis for statements regarding the evolution of art forms in advertising.

VI. COLOR IN MAGAZINE ADVERTISING ³

In order to determine the extent to which color has been used in magazine advertising during the great expansion of advertising during recent years, a statistical study was made as follows:

In four representative magazines—*American*, *Collier's*, *Ladies' Home Journal* and *Country Life*—count was made of the percent of colored advertisements among full page advertisements (color being used only with full page advertisements), exclusive of cover pages.

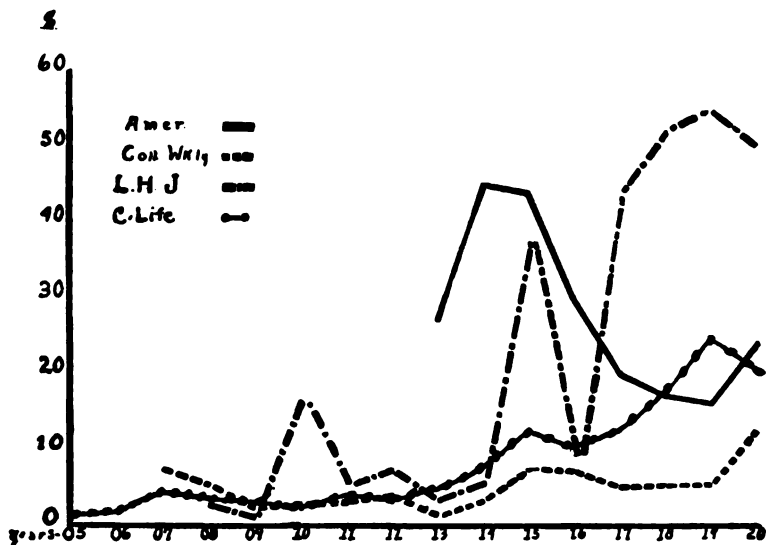


FIG. 1.—Showing per cent of colored advertisements in four magazines, 1905-1920

³ With the assistance of N. Hottel, E. Sample, M. E. Smith, and R. Stegemeier.

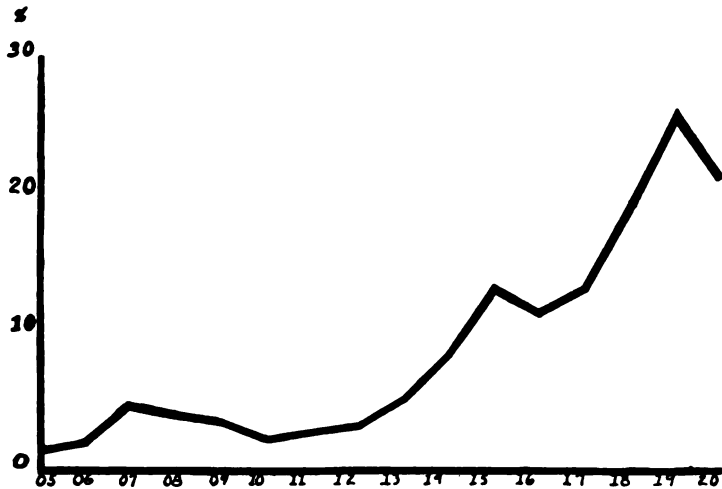


FIG. 2.—Showing per cent of total colored advertisements in four magazines, 1905-1920

RESULTS AND CONCLUSIONS

Results (as shown in Table I and Figures 1 and 2) show that *Collier's* and *New Country Life* were the first of these periodicals to use colored advertisements; the former as early as 1905, the latter in 1907.

After the *Ladies' Home Journal* and *American* entered the field of color in 1908 and 1913 respectively, there was a slight continuous increase until the period of the opening of the World War, when the rise became considerable. The slight decrease of 1916 may reflect a difficulty in securing imported pigments. Recovery was rapid, however, and under the stimulation of post-war business expansion the average amount rose to 24 percent. The decrease of 1920 doubtless reflects the decline of business in general.

Figure 1 shows marked differences between the amounts of color in the several periodicals. *Collier's*, even at its greatest height, has not exceeded a ratio of 12 percent, while the *Ladies' Home Journal* rose to 55 percent in 1919. (Examination of a copy of this magazine for December 1921 showed 89 percent of the full page advertisements to be colored).

The great increase here depicted, most of which has occurred within the past ten years, shows very dramatically the marvelous growth that has occurred in respect to this one of the

ingredients of modern advertising. And it shows that despite the added expense advertisers have found colored advertisements profitable enough to warrant their use almost as frequently as any one of the less expensive art forms discussed in the preceding study of this series.

VII NEGATIVE SUGGESTION IN ADVERTISING ⁴

In applying the laws of suggestion to advertising and selling, psychologists have recommended the abandonment of negative and the use of positive suggestions.⁵ In the effort to see to what extent advertisers, whether wittingly or not, have acted upon this principle, the writer undertook an historical investigation, consisting of the tabulation of the number of advertisements containing negative suggestions, in magazine issues covering a period of twenty years.

A negative suggestion was defined as (a) a phrase containing "no," "none," "nothing," or "not;" (b) the mention of any act on the part of the buyer which would deny the use of the commodity advertised, e.g., "Do not wear yourself out by sweeping;" (c) the mention by word or through illustration of any article competing with the one being advertised.

The advertisements examined were taken from three representative magazines: *Harper's Weekly*, *Literary Digest* and *Collier's*. Six months' issues were examined at intervals of five years, beginning with 1900. All the advertisements, exclusive of those on cover pages, were observed, and count was made of the number containing negative suggestions.

⁴ With the assistance of J. H. Bales, M. M. Shirley, and R. C. Thomas.

⁵ See, for example, W. D. Scott, *Influencing Men in Business*, N. Y., Ronald Press, 1st ed. 1911; 2d ed. 1915, p. 159; and H. D. Kitson, *The Mind of the Buyer*, N. Y., Macmillan, 1921, pp. 159-163.

TABLE I
Showing number of Advertisements Containing Negative Suggestions during period 1900-1920

	Harper's Weekly			Colliers			Literary Digest			Total		
	No.	%		No.	%		No.	%		No.	%	
	of	with	N. S.	of	with	N. S.	of	with	N. S.	of	with	N. S.
	Advs.	N. S.		Advs.	N. S.		Advs.	N. S.		Advs.	N. S.	
1900.....	1183	241	20.4				404	38	9.3	1587	279	17.6
1905.....	980	108	11.0	2196**	143	6.53	1275	58	4.5	4451	309	6.9
1910.....	640	91	14.2	2388	190	8.0	2019	130	6.4	5047	411	8.1
1915.....	239	16	6.7	1565	65	4.16	1989	107	5.4	3793	188	4.9
1920.....	***			1779	58	3.25	2543*	90	3.5	4322	148	3.4

*1919, **1903, ***Discontinued 1916.

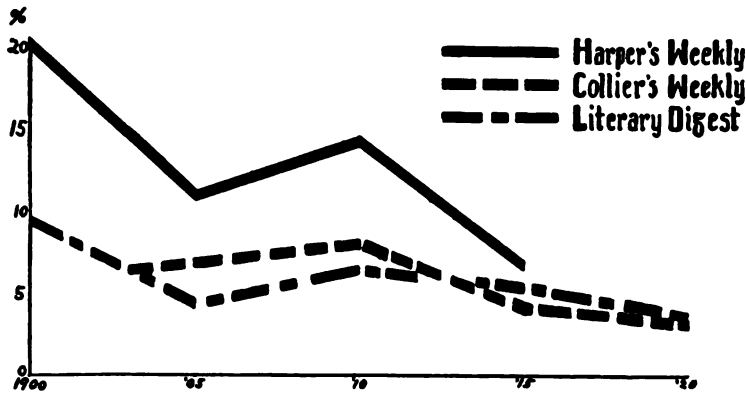


FIG. 1.—Showing per cent of advertisements containing negative suggestions in three magazines, 1900-1920

RESULTS AND CONCLUSIONS

Table I and Figure 1 show that of the advertisements investigated, 17.6 percent on the average contained negative suggestion in 1900. The number has now decreased, however, until there are only about 3.4 percent, or one-fifth as many.

The decline appears in all the periodicals examined, with fluctuations at approximately the same period.

The slight increase in 1910 may be due to the fact that advertising was taking on great vigor about that time and that in their zeal for achieving impressiveness advertisers were tempted to use quantity regardless of the value of their suggestions.

Of the three varieties of negative suggestion mentioned above, that of mentioning in a derogatory way a substituting or competing article was most common.

These results, constituting as they do, a record of the practices of advertisers, give basis for condemning the use of negative suggestions, all apart from theoretical grounds. In so far as historical evidence is indicative of the financial success of a mode of appeal, they indicate that advertisers have found it increasingly profitable to reduce the number of negative suggestions, until today they occur on the average only about 3 times in a hundred.

A SOCIAL RELATIONS TEST

By M. J. REAM, Bureau of Personnel Research, Carnegie Institute of Technology

There are, according to Professor Thorndike,¹ three types of intelligence—conceptual intelligence, mechanical intelligence, and social intelligence. Most of our present day intelligence tests tap the first mentioned type of intelligence but they make no pretense of measuring the last type, which functions in the social relations of life. This latter field is receiving increasing attention in experimental work. The Social Relations test here presented, attempts to throw light on this problem. It is intended to show the range and type of social interests. It is based on the assumption that the man who has a fund of such information, who has mixed with many social groups, is socially inclined and prefers to associate with people rather than spend his time at home or with books. Salesmen by reputation belong to this “social type.” The test was accordingly given to two hundred salesmen. It has proved of some value in predicting the success of new salesmen.

The test consists of fifty items taken from about twenty lines of social diversion. The following are illustrations:

Socially acceptable items:

Games: Which of the following games requires chairs?

London Bridge..... Flying Dutchman..... Three Deep.....

Going to Jerusalem.....

Lodges: In what organization is eleven o'clock of special significance?

Elks..... Odd Fellows..... Masons..... Knights of Columbus.....

Familiar songs: In the song, what follows the phrase, “I’m coming back to you, my Hula Lou”?

“So wait for me”..... “To make you mine”..... “In my canoe”..... “Beside the sea”..... “To old Hawaii”.....

Church songs: In the song, what follows the phrase, “Blest be the tie that binds”?

“Us in thy Kingdom, Lord”..... “My faith on Calvary”..... “Loved ones of kindred minds”..... “Our hearts in Christian love”.....

¹ Thorndike, E. L. Intelligence and its uses. Harper, 140: 227-35. Jan., '20.

Politics: What is a caucus?

A national political convention..... An official county election
..... A meeting of politicians within one party..... A
secret political meeting in violation of law.....

Etiquette: What is "French leave"?

A dance..... Very few odds and ends left over..... Per-
mission easily obtained..... Showing very polite manners
..... Slipping away without notice.....

Parliamentary practice: How may the maker of a motion
withdraw it after it has been put before the house?

By a two-thirds vote..... By a majority vote..... By the
consent of the chairman..... By the consent of the second
.....

Slang: "Fussing" is a slang expression for what?

Fudging..... Nursing a sprain..... Playing the curb mar-
ket..... Having a date.....

Sport items:

Baseball: What is the nickname of the Chicago Nationals?

Cardinals..... Braves..... White Sox..... Cubs.....

Golf: Which of the following golf clubs has a wooden head?

Brassy..... Niblick..... Mashie..... Cleek.....

Prize fighting: What kind of blow is a haymaker?

A hook..... An uppercut..... A broadside..... A jab.....

Horse racing: What is the name of the greatest living race
horse?

Night Mist III..... Man O' War..... Dan Patch..... Sir
Barton.....

Football: How many points are scored by a touchback?

0..... 2..... 3..... 6.....

Bowling: What is the score called when all ten pins are
knocked down with two balls of one inning?

A strike..... A little slam..... A spare..... A break.....

General: What kind of race is a Derby?

Trotting..... Pacing..... Running..... Hurdling.....

Questionable items:

Dice: How many spots on dice make a "Little Joe"?

3..... 4..... 7..... 11.....

Liquor: Which of the following wines is the strongest?

Claret..... Champagne..... Sherry..... Burgundy.....
Bordeaux.....

Cards: What beats a flush?

Fours..... A straight..... Three of a kind..... Two pairs
.....

Theatre: What is the term applied to short, lively chorus
girls?

Kittens..... Ponies..... Baby dolls..... Footlight dodgers
.....

Dancing: Which of the following is best for jazz dancing?

Fox-trot..... Paul Jones..... Two-step..... Waltz.....

March.....

Billiards: What is the name of the shot in billiards in which
the cue ball is struck above the center?

Side shot..... Force shot..... Follow shot..... Draw shot
.....

Roulette: What is usually the highest number on a roulette wheel?
21..... 36..... 40..... 55.....
Stock market: What does "cover on stock" mean?
Take it off the market..... Cash in..... Buy more of the same..... Put up money.....
Slang: What does the expression, "three fingers," stand for?
A glass of whiskey..... A glass of beer..... A pickpocket.....
A weak hand shake.....
General: What kind of news does the Polize Gazette chiefly contain?
Police court news..... Sport news..... Society news.....
News of police organizations..... Abstracts of new laws.....
.....

It will be noticed that the test contains three general classes of items, the socially acceptable, the sport, and the questionable items. While these classes of items appear to be mixed, they are so arranged that each group may be separately scored. The first half of the odd numbered items are socially acceptable, the second half of the odd items are sport and all the even numbered items are questionable. Relative standing on the three classes of items is very illuminating concerning a given individual's interests and habits.

The test follows the usual multiple choice principle of most information tests. The distractors have been carefully chosen to eliminate successful guessing as far as possible. In scoring, each correct answer is given one point. The test is untimed in order to eliminate the large speed factor operating in intelligence tests. Two alternate forms have been provided with fifty items in each form. The test in its original form contained one hundred items. It was divided into two forms on the basis of the difficulty of the items, which was determined by the number of times each item was missed. The resulting two forms showed correlations, on forty-eight cases, of .97 and .96 with the total score.

The Social Relations test was not intended to be an intelligence test, although it resembles the information type of standard intelligence tests. To determine to what extent the Social Relations test differed from the usual intelligence tests, correlations were computed with a special omnibus test adapted from the Army Alpha by the Bureau of Personnel Research.² The correlations on forty-eight cases were:

² An account of this test has been published. Bingham, W. V. Some group tests. *Psychol. Bull.*, February 1920, p. 57.

Social Relations test:	r
Score on socially acceptable items with Spiral omnibus intelligence test71
Score on sport items with Spiral omnibus intelligence test..	.53
Score on questionable items with Spiral omnibus intelligence test38
Total score with Spiral omnibus intelligence test.....	.60

An intelligence factor is undoubtedly present but for inter-test correlations, these are not high.

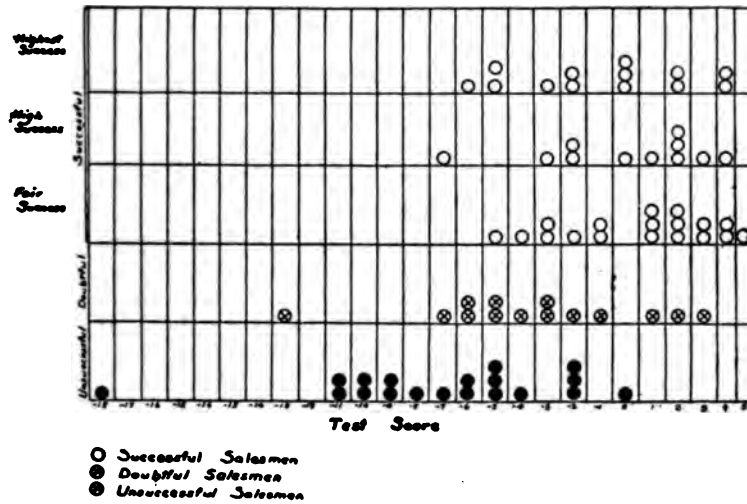
Other correlations were:

	r
Total score, Social Relations test with Education.....	.07
Total score, Social Relations test with hours spent in mechanical work09
Total score, Social Relations test with age.....	-.20

In considering the predictive value of the test for success in sales work, a detailed analysis of the separate items was made. A number of items gave statistical evidence of being significant though the total score on the test gave no such evidence. Older men are more successful salesmen while young men as a rule have higher total scores on the test. But when only the significant items of the test are considered in the scoring, discrimination appears. This is shown in the accompanying chart of insurance salesmen. Success in selling is determined only by amount of sales and number of cases sold while attending a three months sales school.

The test seems to have more negative than positive value for selecting insurance salesmen. In other words, ignorance of social diversions and lack of social contacts is unfavorable to success, but familiarity with many social diversions does not necessarily insure success in selling.

Aside from the test's predictive value, it is a valuable aid in throwing immediate light on an applicant's interests and his activities outside of working hours.



LEGEND FOR CHART

FIG. 1.—Predictive value of Social Relations test for success in selling insurance

X—Composite score on significant items of Social Relations test

Y—Success in selling insurance

THE LETTER OF APPLICATION IN VOCATIONAL SELECTION

By A. T. POFFENBERGER and V. H. VARTANIAN

In our search for tests for purposes of vocational selection, there has been a tendency to turn our backs upon some of the more primitive methods. Among these is the letter of application. As commonly used, it is undoubtedly ineffective. There has been lacking the check upon results so necessary for determining the value of any measure. Letters of application are usually received by one individual who reads them and chooses the one or two that seem to him to represent the desirable applicants. Whether he has chosen the best or whether he has turned the best away, he may never know, because he has only the persons retained as the record of his choice. Perhaps his own judgment is not to be relied upon. His reaction to letters may be influenced by foolish prejudices either temporary or permanent. He might judge differently if he examined the letters again a week later; or the choice might be different if he were to ask the cooperation of ten of his business associates in evaluating the letters, and were to accept the results of their combined judgment. In spite, however, of its shortcomings as commonly used, the letter of application is still considered a part of the routine of the employment process by many institutions and individuals.

Now a letter of application, if really written by the applicant, might show many characteristics that are important and otherwise difficult to measure, such as ability to express oneself clearly and logically, neatness, good taste, seriousness of purpose and care for details. If checked up by some means as to accuracy of statement, it might conceivably serve as a good measure of honesty and integrity.

This report is intended to show to what extent the letter of application might be of service in vocational selection. One other study of similar nature will be briefly reviewed. Walton¹ studied the judgments, by fifty business and professional men, of twenty-five letters of application for a position as bookkeeper. Although there was no check upon

¹ An unpublished Master's thesis in the department of psychology of Columbia University.

the qualifications of the applicants, some interesting results were obtained.

1. *Consistency of the Judges.* Each judge arranged the letters in an order of merit twice, the interval between arrangements being long enough to rule out the factor of memory. The consistency of the judges is shown by the coefficients of correlation² between the first and second arrangement for each judge. The following figures show the average of the coefficients and their range for each of the four traits, intelligence, tact, reliability and neatness:

	Relation between first and second arrangement	Range of Coefficients
Intelligence	.537	.08-.72
Tact	.473	.18-.72
Reliability	.474	.20-.73
Neatness	.600	.14-.91

The figures show that the consistency of judgment varies with the trait; that individuals differ in their consistency in making such judgments; and that there are considerable differences among people in their ability to judge the trait itself from a letter.

2. *Agreement Among the Judges.* The judges agree more closely concerning the presence of some traits than of others when they are sought in letters of application. Those upon which there is most agreement are intelligence and tact, those upon which there is least agreement are reliability and neatness.

3. *Correlation Between Traits.* The coefficients of correlation between traits are given in the following figures:

Intelligence and Tact	.92
Intelligence and Reliability	.91
Intelligence and Neatness	.82
Tact and Reliability	.91
Tact and Neatness	.85
Reliability and Neatness	.84

They support what has frequently been reported, namely, that in such judgments it is not possible to react separately to particular traits, but rather the judge is governed by his general impression. Thorndike has aptly called this the "halo" effect.

4. *Judgment of Good and Poor Letters.* There is closer agreement upon the poor than upon the good letters for three of the traits judged, as shown by the following average devia-

² All coefficients of correlation presented in this report were obtained by the Spearman formula.

tions for the judgments in the case of the first and last letters:

A. D. for the 25th letter	4.02
A. D. for the first letter	5.84

If this were found to hold consistently, it would be advisable to select the best by elimination of the worst rather than by direct selection of the best.

The experiment to be reported here was performed several years ago. It differs from that of Walton in that it provided a check against which the judgments of the letters could be evaluated. Twenty-five students in the senior class of a training school for religious workers in New York City wrote letters of application for a position as religious worker in a certain section of the city. These letters just as received were given in turn to twelve members of the staff of the Union Theological Seminary, New York City, with the request that they arrange them in an order according to the degree to which they indicated general fitness for the position.

The measures which were to serve as checks against the letters were obtained as follows: Five teachers from the staff of the training school furnished three separate arrangements of twenty-five individuals according to the degree to which they possessed the three traits, general ability, intelligence and tact. These three traits were thought by the teachers to cover the qualities most needed for success in the religious work. In addition each member of the group of applicants arranged his twenty-four associates and himself in an order for each of these three traits. The conditions for such estimates were especially good, as the school is small and every teacher has very close association with his students; and the students themselves are more closely associated than are students in the ordinary college.³

We have then for purposes of comparison the following material:

- a—Estimates of general fitness determined from the letters.
- b—Estimates of general ability by the teachers.
- c—Estimates of intelligence by the teachers.
- d—Estimates of tact by the teachers.
- e—Estimates of general ability by the associates.
- f—Estimates of intelligence by the associates.
- g—Estimates of tact by the associates.

³ The average deviations of the positions assigned the individuals ranged from 1 to 5, where the total possible range is twenty-five. The average of these average deviations in the case of intelligence was 3.7. According to the formula presented by Berliner in the *J. of Appl. Psychol.*, 2, 1918, pp. 236ff, the average of all of the deviations for arrangements by chance would be 6.25.

1. *Correlation Between Traits.* When the individuals are judged directly by the teachers and associates, the traits are found to be related as follows:

General ability and intelligence	.95
General ability and tact	.82
Intelligence and tact	.86

These coefficients are about the same as those found by Walton where the letters were judged, and they indicate the tendency to judge from general impression, as previously stated. It would seem that, for practical purposes, a judgment for general fitness would be better than judgments for special traits, as such a task would be much less likely to confuse the judges.

2. *Relation Between the Estimates of Teachers and Associates.* The following coefficients show the relation between the average arrangement of the twenty-five individuals by the teachers and by the associates for the three traits:

	Relation between teachers' and associates' estimates
General ability	.90
Intelligence	.83
Tact	.59

The coefficients for general ability and intelligence are rather high, while that for tact is very much lower. This is, doubtless, due to the fact that the associates and teachers have different criteria of tact. There are no objective measures of this trait to form a common basis for judgment such as school grades would afford for intelligence or general ability.

3. *Estimates of the Letters.* The average deviations for the positions assigned to the individual letters ranged in size from 1.5 to 6.7 positions, where the total range of positions is twenty-five. This range is somewhat greater than in the case of the direct estimation of the individuals for intelligence (1-5). The average of all the deviations for all letters is 4.2 positions, where it would be 6.25 if the arrangements were chance arrangements. (See note above.) This is somewhat higher than the average of the deviations for the direct estimation of individuals for intelligence, namely 3.7. That is, there is a closer agreement in judging the individuals than in judging their letters.

As reported by Walton the agreement among the judges is closer for the poorest than for the best letter (A. D. of 1.5 as compared with A. D. 2.2), but as the difference is small and the relation is reversed for the next poorest and the next best, it has no particular significance.

4. *Relation Between Estimates of Letters for General Fitness and Direct Estimates of the Individuals.* The following figures show the correlations between the estimates of the letters and the estimates of the individuals for the three traits, general ability, intelligence, and tact by teachers and associates.

General ability (Combined estimates of teachers and associates) and letters50
General ability (Teachers' estimates) and letters.....	.56
General ability (Associates' estimates) and letters.....	.46
Intelligence (Combined estimates of teachers and associates) and letters44
Intelligence (Teachers' estimates) and letters58
Intelligence (Associates' estimates) and letters.....	.44
Tact (Combined estimates of teachers and associates) and letters22
Tact (Teachers' estimates) and letters20
Tact (Associates' estimates) and letters18
General ability, intelligence and tact (Combined estimates of teachers and associates) and letters.....	.45
General ability and intelligence (Combined estimates of teachers and associates) and letters.....	.48
General ability and intelligence (Teachers' estimates) and letters55
General ability and intelligence (Associates' estimates) and letters49

A consideration of these coefficients of correlation leads to the following observations:

Coefficients involving tact are the lowest. Walton, however, found rather close agreement among the judges for this trait. It has been noted earlier in this paper that the correlation between teachers' and associates' estimates of this trait was low. On the whole it seems unlikely that letters will provide a good measure of this trait.

Correlations between teachers' estimates and the letter estimates are higher than those between associates' estimates and the letter estimates. Various reasons for this difference suggest themselves, e.g., that the teachers are better judges, or that the two teacher groups (the judges of the letters being teachers) take a more similar attitude toward the task than teachers and associates would be expected to take.

Of the estimates by the teachers, that of general ability and intelligence both correlate fairly high with the average estimates of the letters, namely .56 and .58, in comparison with the usual test correlations. According to Thorndike, a general intelligence test could not be expected to correlate with success in college to a higher degree than about .60.

There is little relation between the estimation of tact from the letters and from the individuals themselves, whether the teachers or the associates are the judges.

If the average positions assigned for the different traits are combined and an order be determined therefrom, the correlation with the order for the letters is not raised above that for one trait alone. The same is true when combinations of two traits are correlated with the letter order.

5. *Relation Between Individual and Group Judgments.* The coefficients of correlation of individual arrangements of the letters with the average order for general ability obtained from the associates are as follows in order of size:

.18, .24, .26, .34, .34, .40, .40, .40, .46, .46, .48, .52

The average of these coefficients is .37. Now the coefficient of correlation of the *average order* for the letters (group judgment) with the average order of the individuals judged by their associates for general ability is .46. Thus the group judgment is somewhat better than the average of the individual judgments. Only two of the individuals show a higher correlation than that of the group. Therefore, although one might find a judge who would do better than the group judgment, the group judgment would be safer unless one had some means of knowing the good judges beforehand.

Now if the same comparisons are made with the teachers' arrangement instead of the associates' arrangement, the following coefficients are obtained, in order of size:

.24, .30, .34, .38, .40, .46, .49, .49, .53, .56, .57

The correlation for the average arrangement of the letters with the teachers' estimates is .56, as compared with the average of the individual coefficients .435.

6. *Consistency of the Judgments of the Letters.* Repeated arrangements of the letters by the same judges were not obtained, because the original letters of application were used and they had become soiled and worn with much handling. Since the conditions for judgment were very much the same as those in Walton's experiment we may assume that the consistency would be about the same.

7. *Basis for Estimation of the Letters.* There were numerous criteria for forming judgments of the letters, such as choice of paper, general style of the letter, handwriting, marginal spacing, line spacing, neatness, and mode of expression. There was also the actual content of the letter, such as the description of previous training, interest in the work, reasons for making application for the position, etc. No effort has been made to determine the importance of these various factors, although it would seem quite worth while to do so. The more mechanical factors such as those first listed could be

eliminated from the judgment, by presenting uniformly type-written copies of the letters. The criteria would then be largely content and mode of expression. In the light of the usual requirement that letters of application be written in the applicant's own handwriting, it would be interesting to know the importance of these factors.

Conclusion. Individual differences in the ability to judge fitness from letters of application must be recognized. Some traits are much more clearly indicated in letters than others, while certain traits like tact can probably not be measured in letters. The combined estimates of a group of judges is more to be relied on than the judgment of one individual unless there is some means of picking the good judge. Some of the correlations here reported are as high as has been expected from the use of intelligence and other more specialized tests as measures of fitness for practical purposes.

The letter of application is not to be relied on even for a preliminary weeding out of candidates unless the ordinary safeguards against personal bias, etc., be used. But if treated and controlled as a test method it should be of service in vocational selection. It might be made to give a composite picture of neatness, intelligence, schooling, truthfulness, interest in details, ambition and many other of the character traits for which at present there are no adequate measures.

NEWS AND COMMENT

NOTES CONCERNING FOREIGN PSYCHOLOGISTS

The appearance has just been made of the first number of the *Journal of the National Institute (British) of Industrial Psychology*, recently noted in these columns. The aims of the *Journal* are editorially stated as follows:

"This *Journal* will describe in non-technical language the methods and results of applying scientific knowledge to the human aspects of industry and commerce.

"Its aim, therefore, forms part of the objects of the Institute which publishes it, namely, the promotion of the mental and bodily health and efficiency of every grade of worker. That excellence of work intimately and inevitably depends on the well-being, interest and suitability of the worker, is now generally recognized.

"The *Journal* will accordingly publish accounts of the research and propaganda work carried out in this wide field not only by the Institute, but by other similar bodies and by individual investigators both at home and abroad. It will contain abstracts of current periodicals and reports, and it will review the more important books on the subjects of industrial psychology and physiology. Contributions to the pages of the *Journal* are invited.

"The *Journal* will be issued gratis to every subscriber of the Institute, and it will appear quarterly or at such intervals as the material deemed worthy of publication allows."

This initial number contains brief, readable statements about the progress of several investigations being pushed in the fields of tin-box manufacture, chocolate packing, hiring of clerical workers. News is also given concerning vocational psychologizing in other countries.

In order for non-members of the Institute to secure the *Journal*, a minimum subscription price of one pound is set. The address of the editor is 329 High Holborn, London, W. C. 1.

There has been developing in Spain an application of psychology which has not been noted in these columns—the *Institut d'Orientació Professional* (Institute of Vocational Guidance) of Barcelona, (Sant Honorat No. 5). The Director is Sr. Ruiz Castella who is assisted by a staff consisting of a statistician, anthropometrist and a psychometrist, the latter being Dr. Mira I. Lopez.

The first three numbers of the *Analys*, the official organ of the Institute, just received, describe investigations relative to the measurement of individuals in relation to various occupations, after the usual patterns.

The latest number of the *Analys* of the Institute of Vocational Guidance at Barcelona announces that the *Institut de Psychologie* of the University of Paris has created a Section of Applied Psychology in charge of M. J. M. Lahy.

A publisher's (Barth, Leipzig) list of brochures containing studies in vocational psychology carried on chiefly under the direction of Drs. Stern and Lipmann includes reports of the application of psychology to street-car driving, typesetting, typewriting, telegraphy, aviation and hair-dressing.

American psychologists working in the field of business psychology are being honored by a visit from Dr. Y. Uyeno, Director of the Institute of Industrial Psychology of Tokyo Imperial University. Dr. Uyeno who was a student of Dr. Matsumoto (Yale) has been giving courses for some years in Psychology of Advertising, and more recently has been engaged in making application of psychological methods to industrial problems, chiefly with the end of increasing the efficiency of workers. He has translated Angell's Psychology into Japanese, and has written several books dealing with mental tests and their applications. On leave of absence this year, he is making a tour around the world, visiting laboratories in America and Europe. H. D. KITSON.

BOOK REVIEWS

ELIZABETH KEMPER ADAMS. *Women Professional Workers*. New York: The Macmillan Company, 1921, pp. 467.

This book, consisting of the fruits of a study made for the Women's Educational and Industrial Union, was "designed; first, for the thoughtful undergraduate who is trying to select her occupation, or, having selected it, trying to see its professional and social relations; second, for the young woman two or three years out of college or school who has drifted into a 'dead-end' occupation or from one occupation to another, and is now trying to get her bearings, and to take a longer view; third, for teachers and administrators in colleges and schools—deans, principals, vocational advisers, and so on—who are seeking a broader basis in fact and outlook in their dealings with young people; fourth, for employers who are increasingly turning to the colleges and professional schools for young men and young women workers; fifth, for men and women everywhere who are considering the scope and nature of the professions and the implications of the participation of women in the worlds of inquiry and affairs."

It takes in some sense the place of the earlier *Vocations for the Trained Woman* (published 1910) superseding the latter by reason of its greater wealth of information and the inclusion of the many new avenues opened to women by the stressful changes of the past decade.

The author would prefer not to treat the professional problems of women separately from those of men, and hopes that the day may come when it will be neither necessary nor profitable so to do. So far as possible, then, she tries to make the treatment of each profession interesting to workers of both sexes; giving scales of wages, kind of training required and illuminating facts about the nature of the work to be done.

Professions treated are Medicine, Law, Ministry, Food and Living Services, Civic and Government Services, Social Services, Personnel Services, Industrial, Commercial and Information Services, Art Services, Technical Services. Psychology is included in the last of these groups, and is accorded sage and generous treatment. The discussion might have been made somewhat more concrete by an investigation of the membership of the American Psychological Association, which in 1921 numbered among its 424 names those of 69 women—16 per cent.

The author notes the well-known fact that women have been regarding the teaching profession with increasing disfavor. In one of her tables showing occupations of graduates of three leading women's colleges, it appears that whereas in 1907 83 per cent entered teaching, in 1918 only 41 per cent entered the profession. The author deplors this fact and advocates that teaching and educational service in the new social setting constitutes the most important single profession for women.

Well conceived in its aim, and practical in its content, the book constitutes a real contribution to the vocational guidance of the growing number of young women who are enrolling in the professional schools of the country and are adding their talent and zeal to the upbuilding of the modern social structure.

A wisely selected and annotated bibliography adds greatly to the value of the book.

H. D. K.

JAMES DREVER. *The Psychology of Industry*. New York: E. P. Dutton & Company, 1921, pp. 148.

BERNARD MUSCIO. *Lectures on Industrial Psychology*. New York: E. P. Dutton & Company, 1920, 2nd ed. rev., pp. 300.

CHARLES S. MYERS. *Mind and Work*. New York: Putnam, 1921, pp. 175.

The appearance of these three books within one twenty-four month period indicates a ripening interest in industrial psychology in the British Isles. They all treat the standard phases of industrial psychology—Vocational Fitness of Workers, Fatigue Study, Time and Motion Study—and all skillfully plead that industry be organized on a scientific basis with an eye to the welfare of the worker.

The slight differences between the three books consist chiefly of individualities in emphasis upon particular points: in the Muscio book arguments designed to meet the objections brought by labor against scientific management; in the Myers book a chapter on Systems of Wage payment; and in the Myers and Drever books the unfoldment of a dream of a central bureau of industrial psychology, similar to the National Institute in England and the Psychological Corporation in the United States. The Drever book contains also two chapters on the applications of psychology in advertising and selling and a series of six tests chosen from the Terman revision of the Binet series which, though accompanied by no norms, are alleged to "yield a result which is not greatly increased in accuracy by the addition of the others."

Upon reading the half dozen or more books dealing directly or indirectly with industrial psychology, the sophisticated psychologist-reader grows wishful that writers would introduce some variation from their unanimous citations of the Munsterberg street-car drivers, the Thompson bicycle ball inspectors, and the Taylor pig iron handlers. Written as they were, however, not primarily for the delectation of psychologists but for the enlightenment of intelligent laymen, these three books should not be expected to bring new contributions to science. The psychologist should be happy to have his case stated so carefully and attractively as it is in these books. And by reason of the experimental work insured by the fostering care of the National Institute of Industrial Psychology there is reason to look to the British Isles for an increasing number of contributions toward the too-scantily clothed body of industrial psychology.

McKINNEY, JAMES, AND SIMONS, A. M. *Success through Vocational Guidance*. Chicago, American School, 1922, pp. 270.

The task of vocational guidance, which is usually regarded as one of measuring the individual at an early period of his life and thereupon designating him for some one vocation, is bound to give way with increasing enlightenment to a saner acknowledgement that the individual has a great number of potential vocational adaptations within him, but that he cannot make any of those that are worth while, without training. The task of vocational guidance thus envisaged becomes that of directing the individual into channels of profitable training and education.

In the work of providing training for men and women seeking better vocational adjustment, which the American School of Correspondence has been carrying on for the past quarter-century, this

institution has been obliged perforce to abandon the theory that a human being is born cut to a vocational pattern. It has helped too many persons change the pattern through the instrumentality of education. It is from such a premise that this book was prepared by the Educational Director and the Director of Foreman Training of the School.

As befits a book designed for the practical, earnest clientele of the School, this volume formulates the problem in downright terms like these: "Every position should be a station on the road of life, not a terminus. Other things being equal, a job is desirable according to the number of roads leading from it to better jobs." "Most of us make many vocational choices. There is no final choice until life ends." Every occupation is thus regarded as a training place where the individual may prepare himself for further advance. Data are presented showing that with each degree of training goes a proportionate measure of advancement. Each occupation described is discussed in terms not only of the kind and degree of training required in it, but also in terms of the pathways which it opens to still higher forms of service.

The leading groups of occupations are treated with as great detail as is possible in the compass of 270 pages; and the reader who has access to the original sources will recognize that the facts are trustworthy.

The book should prove generally useful in the work of vocational guidance in public schools, especially in connection with courses in "Occupations." To the research worker it should prove provocative of investigations designed to supply the great stores of factual material which for the present such books as this are bound to want.

H. D. KITSON.

WILLIAM FREDERICK BOOK, Professor of Educational Psychology and Head of Department of Psychology, Indiana University. *Intelligence of High School Seniors*. The Macmillan Company, January, 1922, 391 pp.

Presents results of state wide mental survey of Indiana High Schools. Develops a practical technique for locating the brightest seniors graduating from high school with a view of providing better educational opportunities for them; determines and presents the range, general level and variations in intelligence of high school seniors going to college and colleges of different kinds, i.e. technical and liberal arts; the intelligence of seniors coming from different economic and occupational classes; the intelligence of those who had been retarded, accelerated, and regularly promoted in high school and the grades; the intelligence of the seniors selecting different high school courses and different occupational careers; the intelligence of seniors representing different communities and schools, etc. The book is divided into three parts: the first describes the method used in making the survey and sets forth why such mental surveys should be made; the second presents the results obtained in this State-wide mental survey. Section three discusses and interprets the results and points out their educational and social significance. It also shows how a mental survey may be used in helping to solve some of the more important practical problems which educators, social workers and progressive business men are called upon to solve, who seek to conserve, cultivate, or utilize the native mental capacities of children and

men. The book develops a method or technique which may be used in making mental surveys of communities and schools. Will be reviewed in later issue of this JOURNAL.

WILLIAM F. BOOK, Head of Department of Psychology, Indiana University. *An Efficient Method for Measuring the Results of Instruction in Colleges and Schools*. Bulletin Indiana University Extension, Vol. VI., No. 12, December 1921.

The report defines and describes an efficient method for measuring school accomplishment, one which will measure the success of a teacher or school and the scholastic achievement of pupils in terms of what may reasonably be expected from individuals with different grades of mental ability. The need for such a pedagogic method is illustrated by materials taken from the results of extensive mental surveys showing marked individual differences in the native mental endowment of the pupils belonging to the same school grades, in the same school system, differences between buildings, school systems taken as a whole, different types of communities, etc. But the chief part of the report is devoted to the results obtained in experimenting with such a method of measuring the scholastic attainment of 300 college students at Indiana University. A technique for making such measurements is worked out and described and the reactions to this more accurate and pedagogic method of reporting the school achievement of students is described and the conclusions drawn from this experiment set forth in detail.

"Personnel Research Agencies" is the title of bulletin No. 299 just issued by the United States Department of Labor, through its Bureau of Labor Statistics (No. 299). It is intended to serve as a guide to organized research in employment management, industrial relations, training, and working conditions.

About a year ago a conference was held in Washington, D. C., under the auspices of the Engineering Foundation and National Research Council to consider the practicability of bringing about cooperation among the many bodies conducting research relating to persons employed in industry and commerce, from management to unskilled labor. This conference, which was attended by 40 representatives of organizations of labor, manufacturers, employment managers, engineers, physicians, educators, economists, and social workers, requested the Bureau of Labor Statistics to undertake a survey of existing agencies whose activities include studies and investigations relating to the employee and his job. The above bulletin contains the results of this survey arranged in the form of a handbook for ready reference, it describes the research activities of about 300 bureaus, associations, foundations, laboratories, and university departments, which deal with the problems of an employment manager's office, vocational psychology, wage systems, cost of living and budgets, employee representation, training of managers, foremen and workmen, whether in educational institutions or in the factory, the relation of hours of labor, fatigue, lighting, ventilation, food, etc., to output and health of workers, occupational diseases and health hazards in the various industries, safety codes and appliances for the prevention of accidents, and the special problems connected with the employment of women and young persons, immigrants, colored workers, the handicapped or disabled, and the mentally deficient or unstable.

The industrial engineer, the employment manager, the trade-union official, the educational director in retail store or factory, the vocational counselor in the public schools, the industrial physician, and the social worker will all find information useful to them in this compendium.

NEW BOOKS RECEIVED

Until further notice books for review should be sent to W. F., Book Department of Psychology, Indiana University.

- H. ADDISON BRUCE. *Self-Development*. A Handbook for the Ambitious. Funk & Wagnalls Company, New York and London, 1921. 332 pp. Price, \$1.50.
- GRACE OWEN. *Nursery School Education*. E. P. Dutton & Co., New York. 176 pp.
- KENNETH RICHMOND. *The Curriculum*. Constable & Company, London, 1919. 144 pp. Price, 5s.
- MARGARET McMILLAN. *The Nursery School*. E. P. Dutton & Company, New York; J. M. Dent & Sons, London and Toronto, 1919. 356 pp.
- JANET PENROSE TREVELYAN (Preface by Mrs. Humphrey Ward). *Evening Play Centres for Children*. E. P. Dutton & Company 183 pp.
- L. A. KALBACH and A. O. NEAL. *Organisation of State Departments of Education*. U. S. Bureau of Education, Bulletin, 1920, No. 46, Washington, D. C. 48 pp.
- . *Survey of the Schools of Wilmington, Delaware*. U. S. Bureau of Education, Bulletin, 1921, No. 2, Washington, D. C. 191 pp.
- The Journal of Educational Method*, Vol. I, No. 3, November, 1921. The World Book Company, Greenwich, Conn. Price, \$3.00 a year, 35c a copy. 128 pp.
- GUY M. WHIPPLE. *The National Intelligence Tests*. The Public School Publishing Company, Bloomington, Illinois. pp. 16-31.
- ARNOLD GESELL. *Hemihypertrophy and Mental Defect*. (Reprinted from the Archives of Neurology and Psychiatry, Oct. 1921, Vol. VI). American Medical Association, Chicago. 24 pp.
- WALTER E. FERNALD and GEORGE M. KLINE. *Massachusetts Department of Mental Diseases*. Vol. IV, Nos. 3 and 4, October, 1920. 135 pp. Boston, Mass. (Papers in Honor of Dr. E. E. Southard's Decennial Anniversary of the Bullard Professorship of Neuro-pathology at Harvard University.)
- DANIEL BELL LEARY. *Philosophy and Education: Part I*. University of Buffalo, 1921. University of Buffalo Studies, Vol. II, No. 2, October, 1921, pp. 45-89.
- SANTE NACCARATI. *The Morphologic Aspect of Intelligence*. Columbia Contributions to Philosophy and Psychology, Vol. XXVII, No. 2, New York, 1921, 44 pp. (Report of Correlations between measurements of intelligence and measurement of extremities and of trunk reciprocally considered and compared.)
- RAYMOND F. CRIST. *Federal Citizenship Textbook*. (A course of instruction for use in the public schools by the candidate for citizenship.) U. S. Department of Labor, Washington, D. C., Part III, 104 pp.
- KENTUCKY EDUCATIONAL COMMISSION. *Public Education in Kentucky*. General Education Board, 61 Broadway, New York 1921, 213 pp.

- GEORGE S. SNODDY and GEORGE E. HYDE. *Mental Survey of Utah Schools and Adaptation of the Army Beta Tests*. Bulletin of the University of Utah, Vol. 12, No. 6, Sept., 1921, pp. 5-28.
- LILLIEN J. MARTIN. *The Work of a Mental Hygiene Clinic for Pre-School Age Children*. Publication No. 5, California Society for Mental Hygiene, San Francisco.
- . *Pedagogical Hints from the Results of a Survey of a San Francisco Public School for Delinquent Boys*. Martin Mental Hygiene Publication No. 5, pp. 5-18.
- WILLIAM F. BOOK. *Preliminary Report on the State-wide Mental Survey of High School Seniors*. Bulletin of the Extension Division, Indiana University, Bloomington, Indiana, pp. 31-67.
- . *Variations in Mental Ability and its Distribution Among the School Population of an Indiana County*. Bulletin of the Extension Division, Vol. IV, No. 4, Indiana University, Bloomington, Indiana, 42 pp.
- PERCIVAL HALL. *Education of the Deaf*. Department of the Interior, Bulletin No. 14, 1921, Washington, D. C., 16 pp.
- . *Present Status of Music Instruction in Colleges and High Schools 1919-20*. Department of the Interior, Bulletin, No. 9, 1921, 53 pp. (Report of a study made under the direction of the United States Bureau of Education by a joint committee of the National Education Association, Music Teachers' National Association, and Music Supervisors' National Conference.)
- GEORGE F. ZOOK and SAMUEL P. CAPEN. *Opportunities for Study at American Graduate Schools*. Department of the Interior, Bulletin, No. 6, 1921, 59 pp.
- SAMUEL PAUL CAPEN. *Facilities for Foreign Students in American Colleges and Universities*. Department of the Interior, Bulletin, No. 39, 1920, 269 pp.

THE JOURNAL OF APPLIED PSYCHOLOGY

VOL. VI

JUNE, 1922

No. 2

A MASS MENTAL TEST FOR USE WITH KINDERGARTEN AND FIRST GRADE CHILDREN

By CLARA HARRISON TOWN, Ph. D.,

Every year finds a greater number of public schools attempting to classify their pupils according to mental ability, and to measure their mental ability by means of mental tests. As an adequate individual mental examination of all children in a large school system is impossible without a large force of mental examiners the mass test has come into quite general use.

There are a number of mass tests adapted for use with children of the third and more advanced grades, but it has been found difficult to devise tests that can be used successfully with groups of kindergarten and first grade children. As the early grades are those in which mental classification is most urgently needed, it is much to be desired that if mass testing in these grades is possible, a satisfactory test be devised.

During the winter of 1919-20 the writer was confronted with the problem of picking out the atypical children in certain kindergarten and first grades in the schools of Kansas City, Missouri, for the purpose of placing them in classes for individual training. Desiring to use for this purpose a mass test an attempt was made to devise one which would bring into play mental activities which Binet found developed in young school children, and which would accomplish this without requiring a verbal or written response from the child.

The writer fully realized the difficulties involved in testing young children in groups, realized the flightiness of attention which renders concerted group work almost impossible, the strong individualistic tendencies, which disappear after several years of class room experience, but which at this period often lead a child to a response quite other than that indicated, the tendency to look to see how others are responding rather than

to depend upon self, which may entirely defeat the purpose of the test, and the short fatigue period, which makes maximum effort throughout a long test series impossible. In spite of this keen realization of the innate difficulties of the situation which made the achievement of a satisfactory test most problematic, the writer thought the demand for such a test justified the expenditure of much time and thought in an effort to devise it.

Bearing in mind the difficult factors in the practical situation, an effort was made to (1) secure test material which would immediately arouse interest and pleasure, (2) avoid confusion of impression by presenting each problem in isolation and (3) guard against wandering of the attention by using a series of tests each of which is quickly solvable and through which the examiner may personally lead the class without delay. It was decided to adapt, wherever possible, the material of the Binet-Simon Intelligence Scale to mass methods. As verbal and written responses were both ruled out, the first by the mass nature of the test and the second by the inability of the children to write, the "cross out" method, so successfully used by Dr. Pressey and others, was linked with the Binet problems and a "Picture Game" evolved.

The "Picture Game" consists of a sixteen page picture book, five inches by seven inches, which is here reproduced. Almost all children love pictures and when one proposes a game to a class and distributes little books full of pictures with which to play it, there is usually a glad and happy response.—interest and pleasure are at once aroused. The success of the test depends largely upon whether this pleasurable interest can be sustained throughout the performance. This the writer tried to insure by a series of short and varied tests which a class can be put through in from ten to fifteen minutes.

The game consists of thirteen tests and an initial page used for teaching what "cross out" means. Tests 1, 2, 3, 4, 5, 6, 7, 9, 11, 12, and 13, in order of presentation in book are adaptations of tests 30, 7, 13, 15, 12, 19, 24, 25, 33, 17, and 23 of the Binet-Simon Intelligence Scale, 1911 Revision (Town translation). These tests in order of presentation are: color naming; pointing out of eyes, nose and mouth; pointing out of objects named; discrimination of length; discrimination of sex; counting four objects; counting thirteen objects; aesthetic judgment; completion test; copying of square; copying of diamond. Two additional tests are included in series—number eight, a form test, and number ten, a completion test. Both of these are based upon well known tests.

In each of the cross out tests, and all are cross out tests with the exception of the drawing of the square and the diamond, three chances are given child to demonstrate ability to deal with the problem. For each test three pictures are presented on one page, but distinctly separated, in all of which the fundamental problem involved in the cross out is identical. If the child succeeds in crossing out two correctly it is assumed that he understands the problem and full credit for test is awarded. Exceptions to this rule are made in the color naming test and in the aesthetic judgment test, in both of which a perfect performance is required. This variation is necessary in the aesthetic judgment test because it is impossible to arrange the three pairs of pictures in such a way that two ugly ones will not be in line with each other, and a child might make two successes by the simple procedure of marking all the pictures on the right or left of the page. The color test also gives much opportunity for chance successes and it was therefore thought better to require three correct cross outs.

Partial credits are used only for drawings of square and diamond. The individual differences in the drawings are so great that a credit of success or failure seems inadequate to meet the situation. The partial credits are based on examiners judgment of (1) general shape of drawing, (2) the accuracy of the angles. Only one attempt at each drawing is allowed.

The performance is scored by means of a point scale. The number of points assigned for each of the thirteen tests varies with (1) the difficulty of the test, (2) the mental age at which ability to perform the test is usually developed. The degree of difficulty was experimentally determined by trying out the tests on 419 kindergarten children and 778 first grade children in several of the schools of Kansas City, Missouri. The number of children failing in each test was recorded and the tests then arranged in rank order determined by the number of children failing in each. The mental age was assumed to be that determined by Binet for the performance of similar tests in the Intelligence Scale. Tests number eight and ten, having no equivalent in the Binet series, were scored entirely on a basis of rank order.

The lowest score assigned to any test is six, this is given to tests placed at the four year level by Binet and the other tests are scored on the principle of one additional point for each additional year of assigned mental level. The rank order of difficulty made it necessary to score the aesthetic judgment test one point lower, and diamond drawing one point higher

than the Binet age ranking would indicate. These are the only conflicts between Binet ranking and our attained rank order of difficulty in spite of the fact that mass instead of individual methods were used. In scoring drawing of square, four is given for shape and three for four correct angles. In scoring drawing of diamond, five is given for shape and four for four correct angles.

The rank order (based on number of failures in each) of the thirteen tests, with scores assigned to each is as follows. The tests are numbered according to their order of presentation in the book, which was determined partially by their supposed difficulty, partially by the desire to have the blank pages for drawings at the end of book and test series, and partially for the printer's convenience in arranging the color pages. It is probably of some practical advantage in giving the tests to have the more difficult ones distributed through the series rather than grouped.

Order of difficulty	Test No.	Test Score	Page
1	4	6	6
2	3	6	5
3	2	6	4
4	5	6	7
5	12	7	15
6	9	7	11
7	6	7	8
8	7	8	9
9	10	9	12
10	1	9	2, 3, 14
11	13	9	16
12	11	10	13
13	8	10	10

PROCEDURE

As indicated by the name of the test series "Picture Game," this series is a game and should be conducted in the spirit of one. No notion of a test should be allowed to enter the children's minds. The children should be seated at as great distances from one another as possible in order to prevent copying. This is particularly necessary with kindergarten children. Not more than twenty kindergarten children should be handled in one group. The tops of desks should be cleared and each child provided with a black crayola.

When all is ready the examiner asks the children whether they would like to play a game, and when they assent she tells them that each will have a little picture book to play

it with. She then distributes the books telling the children to place them on their desks so that the rabbit is looking at them, and to leave them on their desks during the whole game. All is then ready and the examiner proceeds as follows: "Do you see the rabbit? Do you see the rat? Cross out the rabbit. When I say 'cross out' I mean, put a line through it like this." Demonstrate on blackboard and on sample book which you hold up before the class. After rabbit has been crossed out say, "Do you see the duck? Do you see the picture of the book? Cross out the picture of the book." This first page of the book is not a test but is used for instruction in crossing out. The test series is not begun until the examiner is assured by actual inspection of each child's book that all the children know what is meant by "cross out." This is the only bit of knowledge which is required for the manual execution of all the cross out tests. When it is determined that each child knows what is meant by "cross out" proceed as follows, allowing ample time for crossing out between each instruction: "This little book is full of pictures, and on each picture I am going to ask you to cross out something. You cross out just what I tell you to, and do not cross out anything else. Now we will begin. Are you all ready? Turn the page.

Test 1, page 2—"Cross out the red. Page 3—"On the next page cross out the blue. Turn the page.

Test 2, page 4—"See the pretty little girl at the top of the page? Cross out her nose.

See the little girl just below her? Cross out her eyes.

See the last little girl? Cross out her mouth.

Test 3, page 5—"At the top of the next page you will see a picture of a room. Do you see it? Cross out the bed.

Below that picture you will see another picture. Cross out the boat.

At the bottom of the page you will see another picture. Cross out the key. Turn the page.

Test 4, page 6—"There are two pencils at the top of this page. Everyone see them? Cross out the longer pencil, the longest pencil.

Below the pencil there are two rulers. Cross out the longer ruler, the longest ruler.

At the bottom of the page there are two knives. Cross out the longer knife, the longest knife.

Test 5, page 7—"At the top of the next page you see a little boy and a little girl. Cross out the boy.

In the next picture cross out the man.

In the next picture cross out the girl.

Turn the page.

Test 6, page 8—"At the top of the next page you will see some chains. Do you see the chains? Cross out the chain that has four links.

Below the chains there are some apples. Do you see the apples? Cross out the pile of four apples.

Below the apples there are some pears. Cross out the pile of four pears.

Test 7, page 9—"At the top of the next page you will see some acorns. Do you see them? Cross out the line of thirteen acorns.

In the next picture you see a lot of little ducks. Cross out the line of thirteen ducks.

Below the ducks there is a picture of blocks. Cross out the line of thirteen blocks.

Turn the page.

Test 8, page 10—"At the top of the page you see four blocks—a circle, a triangle, a cross, and a square. There are heavy lines under them and just below five blocks. Four of these are like the blocks at the top but there is one left over, which is not like any at the top of the page. Cross out this one that is left over, the one not like any block at the top of the page.

Below this picture is another picture of five blocks. Four of these are just like the four at the top of the page, and there is one left over.

Cross out the one that is left over.

At the bottom of the page there is another picture of five blocks, four of these are just like the four blocks at the top of the page, and there is one left over. Cross it out.

Test 9, page 11—"At the top of the next page you will see two ladies, a pretty lady and an ugly lady. Cross out the ugly lady.

In the next picture you see two horses. Cross out the ugly horse.

In the next picture you see two men. Cross out the ugly man. Turn the page.

Test 10, page 12—"At the top of the next page there are pictures of three objects. Two of them are something alike and one is different. It is not like the other two. Cross it out. Below this picture there is a picture of three people. Two of them are alike. The one is not like the other two. Cross it out.

At the bottom of the page you see the picture of three animals. Two of them are alike. One is not like the other two. Cross it out.

Test 11, page 13—"At the top of the next page there is a picture of a cow. Do you see it? When they drew that cow they did not quite finish her, they left something out. Put a mark where it should be.

Just below the cow there is a picture of a boy's head. When they drew that head they did not quite finish it, they left something out. Put a mark where it should be.

At the bottom of the page there is a picture of a house. When they drew that house they did not quite finish it, they left something out. Put a mark where it should be.

Turn the page.

Test 1, page 14—"Cross out the green.

Test 12, page 15—"You see the next page is an empty page. I am going to draw something on the blackboard, and I want you to draw it on this empty page, just as nicely as you know how." (Draw a square on blackboard.)

"Turn the page.

Test 13, page 16—"The next page is empty too. Now I am going to draw something else on the board, and I want you to draw it on this empty page just as nicely as you know how." (Draw diamond on board.)

"Turn your books over so that the rabbits will be looking at you again.

Write your name on the line marked name."

After name is written, say—"Now write how old you are."

If children have not learned to write, have teacher write names and ages.

EXPERIMENTAL RESULTS

Twelve kindergarten classes, containing in all 419 children, (average age, 5.5) and twenty-two first grades, containing in all 778 children (average age 6.9) were tested by the writer. The number of individuals failing in each test was counted and recorded in terms of percentage of grade group. The average total score, with its mean variation and probable error, and the median, were computed for each separate class and each grade group. Finally the individual total scores of children in each grade group were thrown into percentile tables in order to show what scores the 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 percentiles of each grade group secured. All of these results are here presented in tabulated form.

A group of second grade classes were also examined, but the tests proved to be so far below the level of ability of the second grade children that it is not worth while to present the results.

The results here given confirm the impression gained by the writer in conducting the tests that the Picture Game is adapted to use with kindergarten and first grade children, and that by its use one can obtain a fairly good classification of children in these two lower grades.

PERCENTAGE OF CHILDREN FAILING IN EACH TEST			
Rank order of test	Test No.	Kind. Child. failing—per cent	First Grade Child. failing—per cent
1	4	13	1
2	3	16	2.4
3	2	28	7.8
4	5	36	11
5	12	36	20
6	9	42	15
7	6	52	13
8	7	62	15
9	10	52	23
10	1	57	25
11	13	64	41
12	11	73	43
13	8	75	45

The rank of tests 4, 3, 2, and 5, the four lowest in rank and of tests 13, 11, and 8, the three highest in rank and of test 9 are identical in kindergarten and first grade groups. The differences of rank order are caused by tests 6 and 7, the two counting tests, and are probably due to the facts that count-

CLASS AND GROUP SCORES—KINDERGARTEN						
Class	Numb. Child.	Ave. Age	Ave. Score	M. V.	P. E.	Median
1	41	5.6	54.2	22.9	19.35	62
2	15	5.9	62	18.1	15.29	65
3	47	5.1	41.7	16.5	13.94	42
4	30	5.7	38.5	20.2	17.07	45
5	37	5.7	37.5	19.5	16.48	38
6	16	5.6	41.4	14.9	12.59	34
7	58	5.1	47.3	17.4	14.70	44
8	39	5.6	36.9	21	17.75	34
9	34	5.9	51.7	22.2	18.76	56.5
10	44	5.4	56.9	20.8	17.58	58.5
11	45	5.2	56.4	18.6	15.72	58
12	13		42.1	17.7	14.96	44
Totals	419	60.8	566.6	229.8	194.19	581
Averages		5.5	47.2	19.1	16.18	48.4

MENTAL TESTS FOR KINDERGARTEN AND FIRST GRADE 97

ing is still a novel process for the kindergartners and has already become almost automatic with the first grade children. Test 12 is displaced in the first grade series by tests six and seven which take precedence. The adopted rank order was obtained by pooling the results obtained from kindergartens and first grades. This order is used in the above table.

PERCENTILE SCORES—KINDERGARTEN

Percentiles	Scores
100.....	100
90.....	81
80.....	70
70.....	62
60.....	54
50.....	47
40.....	41
30.....	33
20.....	24
10.....	12
00.....	00

CLASS AND GROUP SCORES—FIRST GRADE

Class	Numb. Child.	Ave. Age	Ave. Score	M. V.	P. E.	Median
1	18	6.9	81.5	9.5	8.03	83.5
2	38	6.3	66.8	17.6	14.87	69
3	40	7.7	70.5	14.9	12.59	75
4	22	8.2	83.8	8.9	7.52	85
5	33	7.2	73.9	12.9	10.90	79
6	32	7.2	57.9	14.5	12.25	56
7	44	7.7	79.8	9.7	8.19	80
8	40	6.6	63.7	13.7	11.58	64.5
9	41	6.9	61.2	11.3	9.55	64
10	23	7.4	88.9	7.2	6.08	90
11	39	6.5	82.1	10.6	8.96	84
12	42	6.3	73.8	14.2	12.00	77.5
13	41	6.6	78.1	14.6	12.34	82
14	42	6.6	78.1	13.8	11.66	82.5
15	28	7.3	72.7	13.5	11.41	79
16	38	8	75.1	12.9	10.90	75
17	39	6.5	76.1	14.7	12.42	80
18	45	7	84.3	11.3	9.55	90
19	40	7.1	81.6	12	10.14	84.5
20	24	6.5	75.2	12	10.14	77
21	36	7.4	72.8	15.5	13.10	81.5
22	33	6.2	59.7	15.3	12.93	65
Totals	778	154.1	1637.6	280.6	2371.1	1637.6
Averages		7	74.4	12.7	10.77	77.4

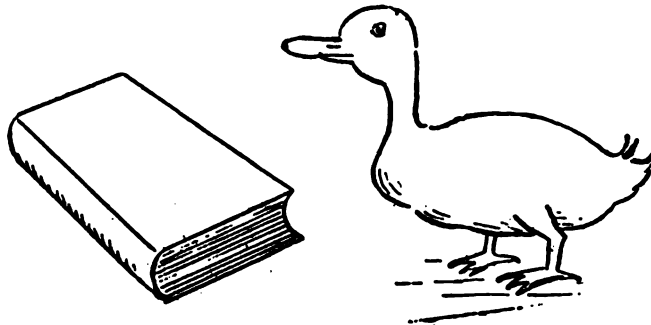
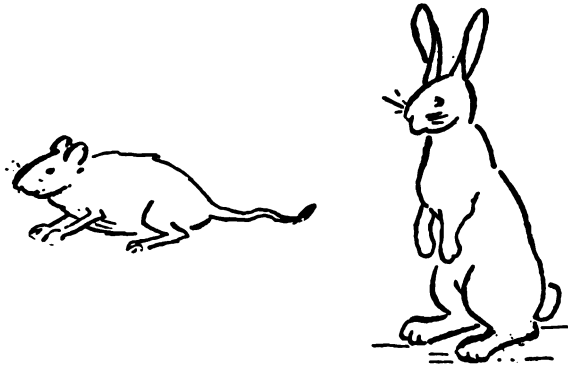
PERCENTILE SCORES—FIRST GRADE

Percentiles	Scores
100.....	100
90.....	96
80.....	90
70.....	87
60.....	82
50.....	78
40.....	72
30.....	67
20.....	59
10.....	49
00.....	8

The tests were also used with a number of Italian and a number of colored classes. The results obtained from both differed so widely from those here reported that the writer omitted them from the compilation and reserved them for future report. A study of the tables here presented reveals in several instances great variation in performance between parallel grades in different schools. A study of the causes of these differences is of interest from many angles, and brings to light many factors of practical importance for these Kansas City Schools. The present article is designed simply to present the Picture Game as a new and possibly useful form of test, to describe it and give explicit directions for its use. To preserve the clarity of this purpose discussion of the problems mentioned in this paragraph is here omitted.

PICTURE GAME

Name _____






RED



BLUE



GREEN



YELLOW

NOTE.—On test sheets blocked spaces are colored as indicated.

MENTAL TESTS FOR KINDERGARTEN AND FIRST GRADE 101

RED

BLUE

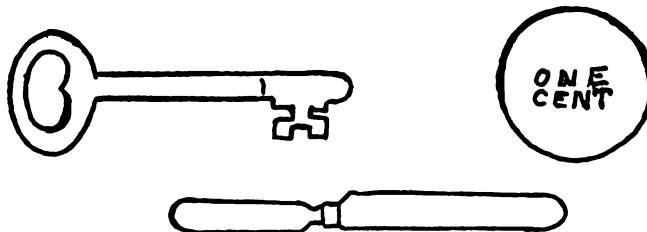
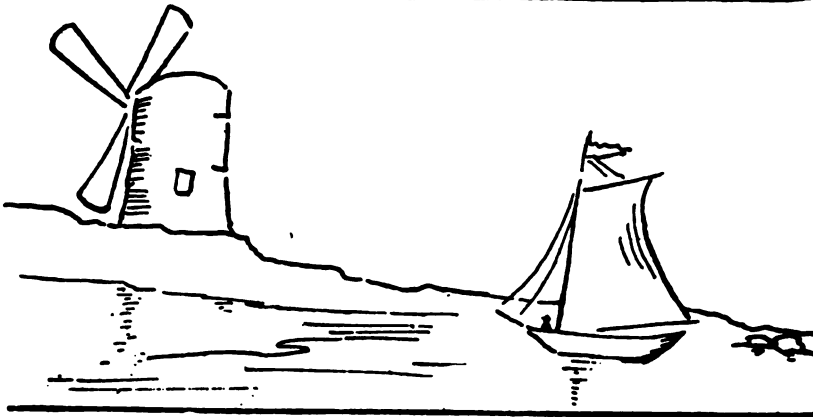
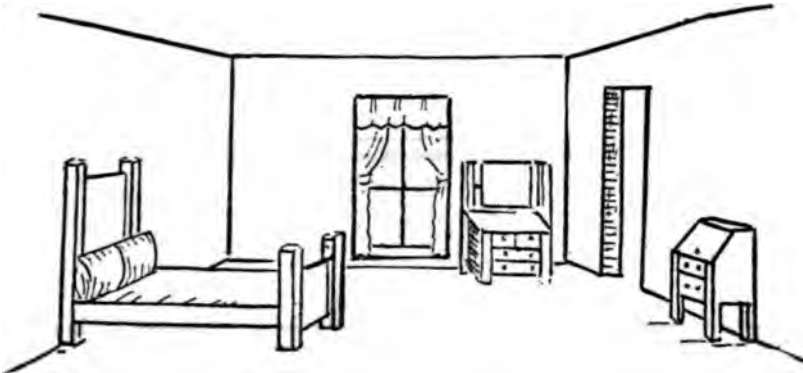
GREEN

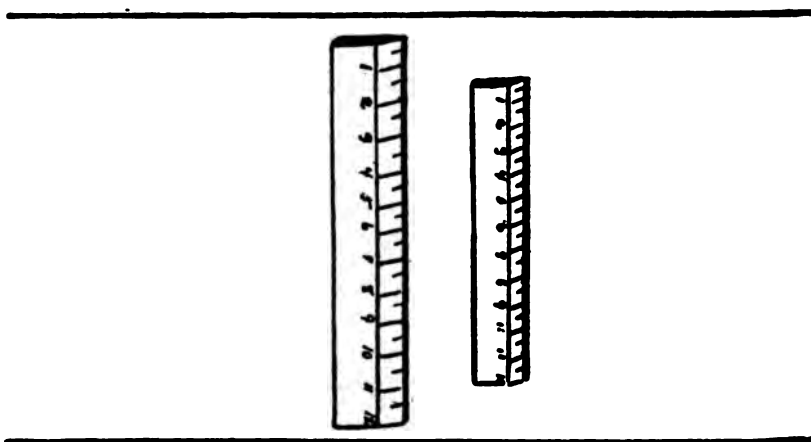
YELLOW

NOTE.—On test sheets blocked spaces are colored as indicated.

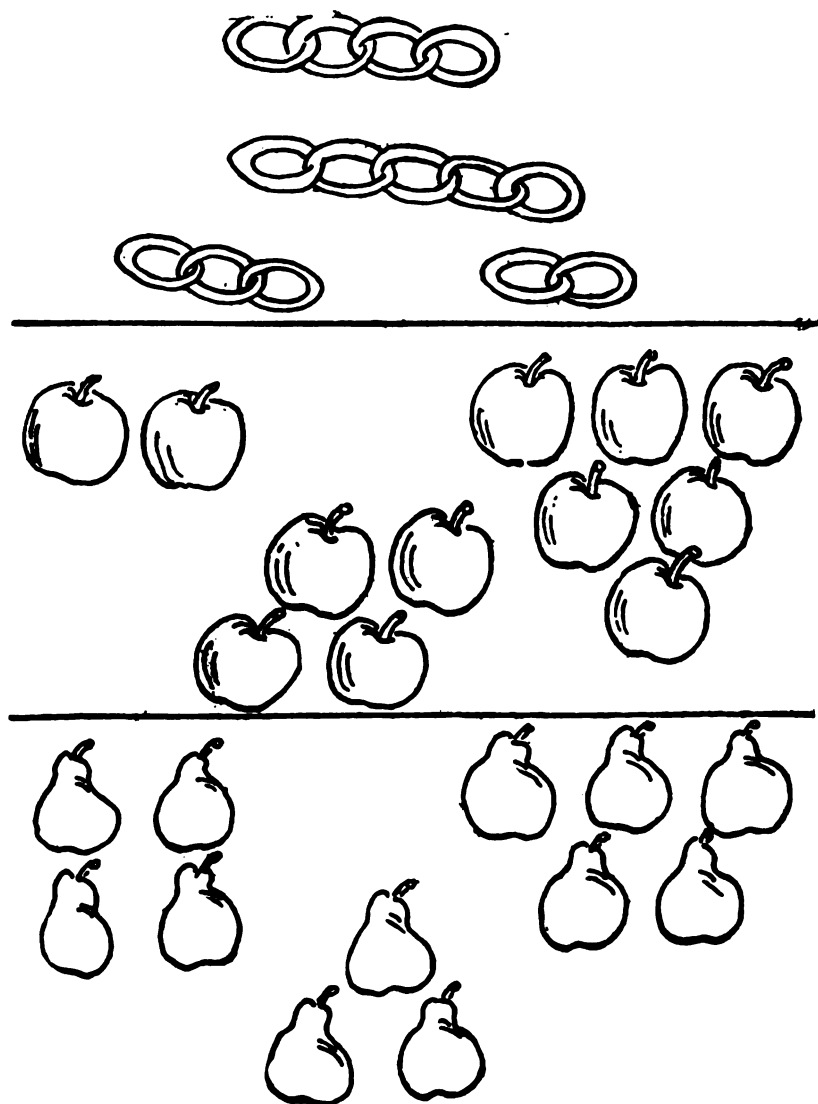


MENTAL TESTS FOR KINDERGARTEN AND FIRST GRADE 103

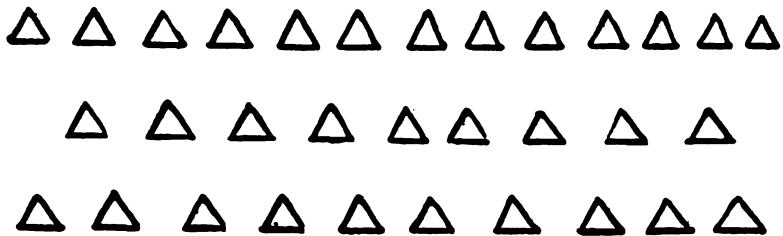
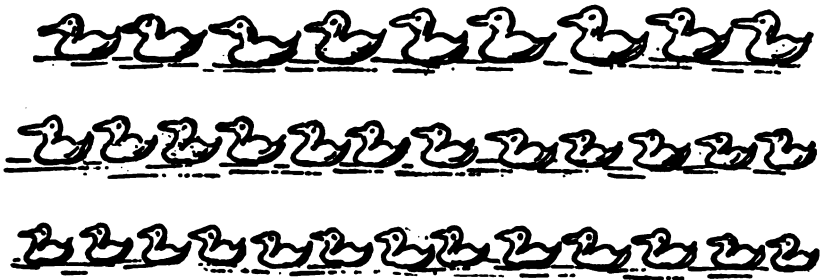


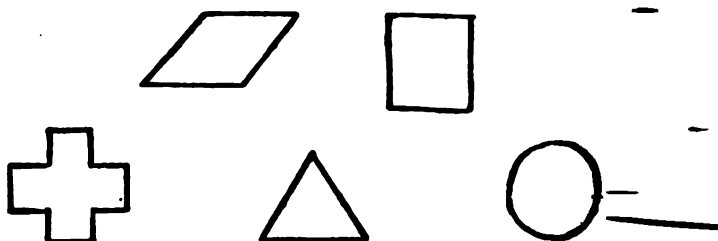
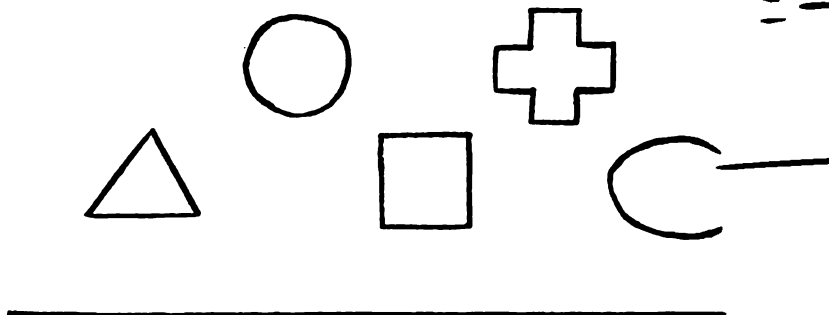
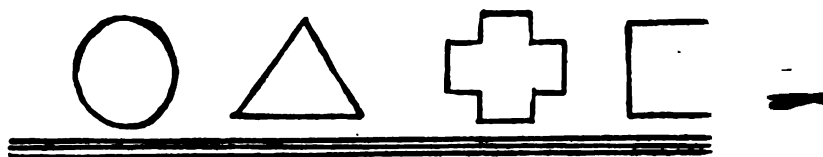






MENTAL TESTS FOR KINDERGARTEN AND FIRST GRADE 107





GRADE 111

11

11
11
11
11

11

11

11

11

11

11

11

11

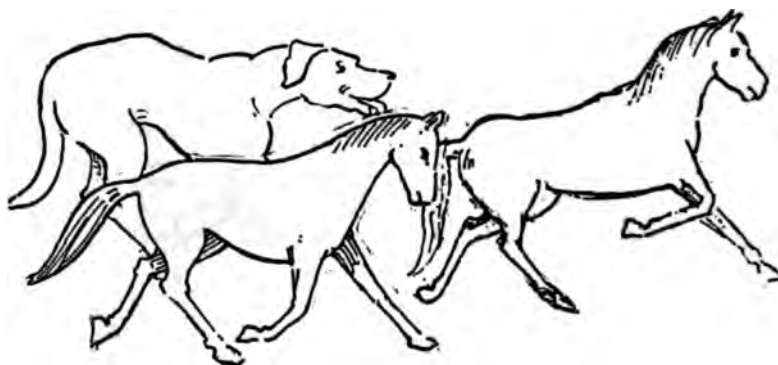
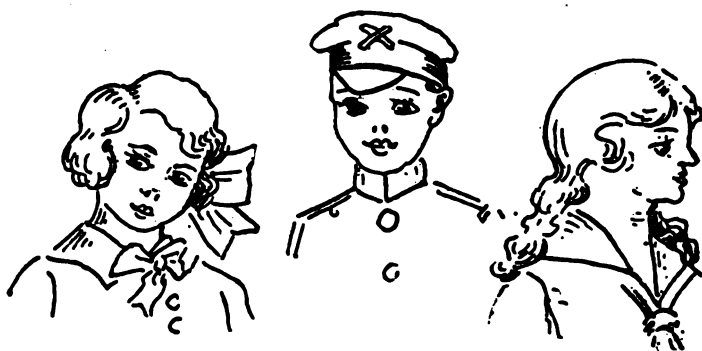
11



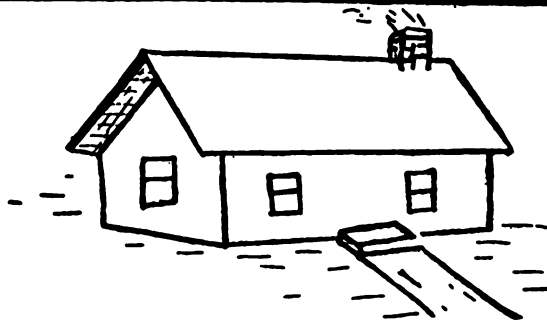
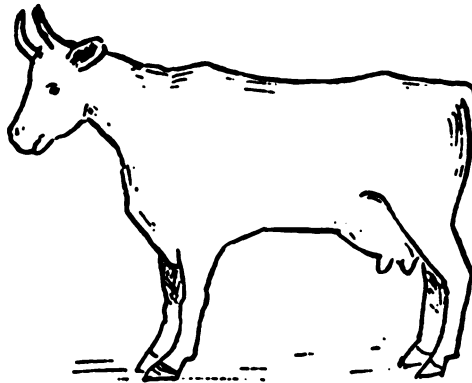
11

11
11
11

11



MENTAL TESTS FOR KINDERGARTEN AND FIRST GRADE 111







RED



BLUE



GREEN



YELLOW

NOTE.—On test sheets blocked spaces are colored as indicated.

The booklets referred to above may be obtained from Clara Harrison Town, 2328 N. 22nd St., Philadelphia, Pa.

STANDARDIZATION OF THE WHIPPLE-HEALY TAPPING TEST

Prepared under the Direction of AMY HEWES, by ELSIE BONITZ, AGNES
DAVIS, PERSIS MOORE, CHARLOTTE PORTER, FLORENCE TURNER,
Students in the Course in Statistics at Mount Holyoke
College

The present study is a standardization of a test for measuring speed and accuracy of eye-hand co-ordinations. It was made in the spring of 1921 by five students in the course in Statistics at Mount Holyoke College. The central tendency was estimated for each age group in a total of 2,253 children, ranging in age from seven to seventeen years, in order that

TABLE 1
NUMBER OF CHILDREN TESTED BY AGE GROUPS

Age in Years	Number			Chicago and Boston	Holyoke and South Hadley
	All Records Total	Boys	Girls		
Total	2253	1444	809	1855	398
7 and less than 8	155	81	74	21	134
8 and less than 9	156	84	72	48	108
9 and less than 10	166	97	69	103	63
10 and less than 11	159	103	56	103	56
11 and less than 12	158	119	39	126	32
12 and less than 13	180	136	44	177	3
13 and less than 14	208	141	67	206	2
14 and less than 15	322	224	98	322	..
15 and less than 16	331	231	100	331	..
16 and less than 17	317	190	127	317	..
17 and over	101	38	63	101	..

the performances in any given case might be rated and used as one of a group of results on which to base vocational advice.

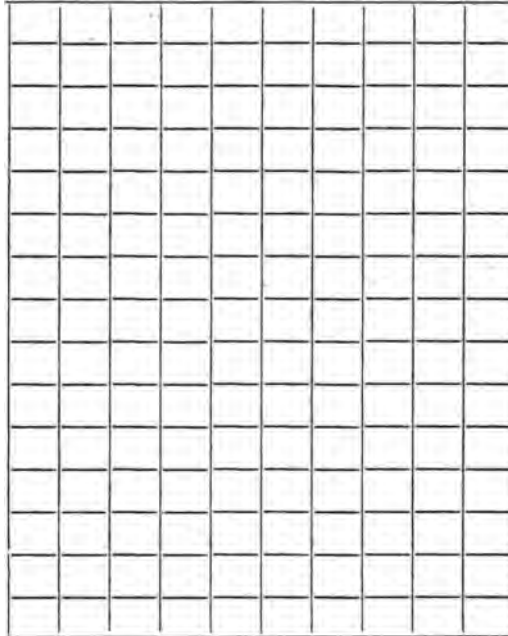
The data comprised the records for 1,855 tests made in connection with the work of the Juvenile Psychopathic Institute in Chicago, and the Judge Baker Foundation in Boston, in both places under the direction of Dr. William Healy and Dr. Augusta F. Bronner, and the records for 398 tests made of

children in schools in Holyoke and South Hadley for the purposes of supplementing certain of the age groups where the numbers were not sufficient. Care was taken to give the latter in schools in different sections of the city in order to get as varied a selection of children as possible. The records were all of children classified as normal mentally.

It was at first thought necessary to secure a minimum of 150 cases in each group, but later, when it was desired to make a comparison between girls and boys, some experimentation led to the belief that 80 were sufficient for a fair basis, for the reason that the median for this number remained unchanged when the group was increased.

THE TEST

The purpose of the tapping test is to determine speed and accuracy of eye-hand co-ordinations of the child. For the test a sheet of paper is used on which is printed a rectangular form divided into 150 half inch squares, as shown in the following diagram.



STANDARDIZATION OF WHIPPLE-HEALY TAPPING TEST 115

The test is given in a room where there is nothing to disturb or distract the child, and the examiner tries as nearly as possible to put him at his ease. The subject is seated at a table and told to tap across the paper (placed as illustrated) once in each square, as rapidly as possible without making any mistakes, the procedure first being demonstrated. An error consists of putting more than one dot in a square, touching the line, or skipping a square. Thirty seconds are allowed for each trial. After the first trial the paper is turned around and the test is repeated.

The best of these two records is used in grading.

DIRECTIONS

The examiner says, "This is a test to see how well you can use your hands. This is how you do it. When I say, 'Go,' you are to put a little tap like this (show how the taps are rapidly made across one row and back on the next line), one in each square, never missing a square and not touching any lines. See in how many squares you can tap till I say 'Stop.' Sit comfortably with your arm the way you want it."

Have a stop-watch in hand and be in position to closely observe performance for errors. Start watch exactly as you say "Go," and say, "Stop," when watch is exactly at 30", noting where pencil of subject is at this time.

After first trial turn paper upside down, erasing dots if space is needed, and get second record for 30" on the same sheet, saying, "Now, let us see if you can do better a second time."

If the subject is too careful about making large dots or lines warn him that only light taps are necessary. Endeavor to stimulate if lethargy is shown on first trial. If hesitation because of error at any time, tell subject quickly to go ahead.

SCORING

In scoring the test three points are taken into consideration: the speed, which is expressed numerically as the total number of squares tapped; the error, or the number of mistakes which are made; and the index, which is arbitrarily defined to be the speed minus twice the number of errors.

In establishing standards for the different age groups the median was chosen to indicate the normal performance of the children of a given age. Since one-half of the children stand above this point and one-half below, the figure may be taken

STANDARDIZATION OF WHIPPLE-HEALY TAPPING TEST 117

as descriptive of the ability of the ordinary or normal child, and is therefore called a norm.

The individual performances of the children of the different age groups are exhibited in Table II. The figures show a marked and continuous development in motor co-ordination as the age increases, with only a slightly smaller increase in the rate of improvement for the older children than for the younger.

TABLE III
MEDIAN FOR TAPPING TEST

Ages of Children Tested	Total	Speed Boys	Girls	Medians Error			Total	Index	
				Boys	Girls	Total		Boys	Girls
7 and less than 8	45	44	47	0	0	0	43	43	45
8 and less than 9	52	50	53	1	1	1	48	46	53
9 and less than 10	58	56	60	1	2	1	55	52	58
10 and less than 11	65	62	70	1	1	1	62	61	69
11 and less than 12	70	67	76	1	1	0	67	65	75
12 and less than 13	72	71	77	1	1	1	70	69	74
13 and less than 14	76	75	81	1	1	0	74	72	80
14 and less than 15	82	80	86	1	1	1	79	77	84
15 and less than 16	84	82	90	1	1	1	82	80	88
16 and less than 17	86	85	89	1	1	1	84	82	87
17 and over	91	93	90	1	1	0	89	90	89

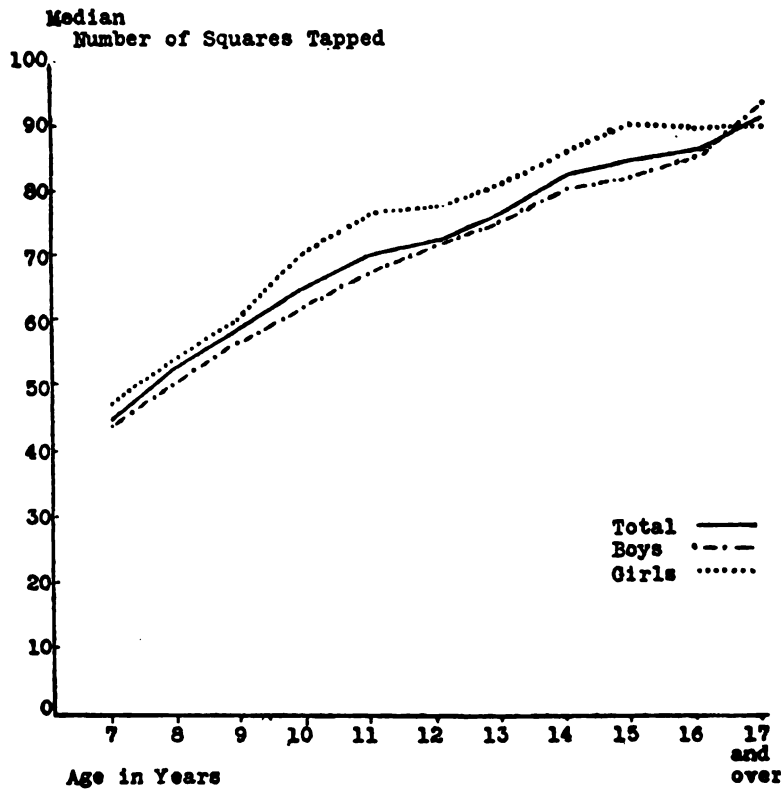
It will be observed from Table III that the speed with which the task is accomplished is the important factor in determining the index to represent the performances of the groups of children of various ages, for the number of errors is small and appears to have slight relation to age. One error was the median for all ages except for the youngest children. If the boys and girls are taken separately, it is seen that the majority of the girls in three of the groups of older children made no errors. As a consequence of the small number of errors, the index number follows very closely the general trend of the curve for speed.

A question which is opened by these figures is that of the relative ability of boys and girls. The curves show that the girls tested have developed motor skill earlier than boys, and that they maintain a small, but nearly constant degree of superiority until the age of 17, approximately, is reached, when they are overtaken by the boys. The numbers tested are not sufficiently large to warrant generalizations as to the relative ability of girls and boys of various ages in other performances of this kind. They do, however, indicate that the field would be a fruitful one for further experimentation, especially in view of the importance of the results for practical use.

A.D. Jones II

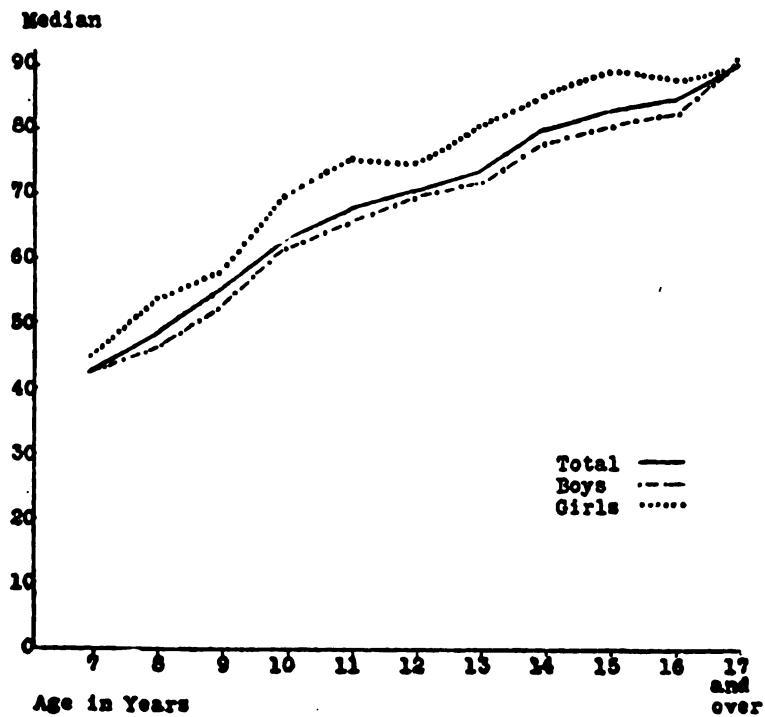
old
p/c
36

FIGURE A
MEDIANS FOR SPEED BY AGE GROUPS



STANDARDIZATION OF WHIPPLE-HEALY TAPPING TEST 119

FIGURE B
MEDIANS FOR THE INDEX BY AGE GROUPS



SOME FACTORS OF AESTHETIC JUDGMENT

By MRS. ANNE ROSS REYMERT

CONTINUED FROM MARCH ISSUE

EXPERIMENT XVII¹

Series S. Judgment based on degree of suggestiveness of an idea,—i.e., the concept of "Spring" in illustrated Spring advertisements of similar size, non-colored, from which all word matter was removed. The advertisements involved a great variety of suggestions, which fact was fruitfully brought out by the reasons for choice noted by the observers.

From the average rankings, which give an idea of the placement of the pictures:

Labels of pictures	Girls	Boys
1	3	1
2	1	5
3	2	6
4	9	7
5	10	10
6	8	4
7	6	3
8	5	8
9	7	9
10	4	2

the following individuals correlations with the average group rankings of the sex-groups, respectively, are derived:

GIRLS:												
mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs	
0.89	0.36	0.64	0.79	0.79	0.27	0.59	0.20	0.24	-0.09	0.67	0.90	
BOYS:												
CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK	
0.07	0.52	0.50	0.03	0.27	0.85	0.49	0.28	-0.52	0.48	0.38	0.48	

Average correlation of girls as a group 0.52

Average deviation of girls as a group ± 0.27

Average correlation of boys as a group 0.32

Average deviation of boys as a group ± 0.24

The correlation between girls and boys as one group=0.50.

¹ In all the following experiments the observers were requested to note the reasons upon which their judgments were based.

That the girls as a group agree so much better among themselves, than do the boys among themselves, seems to be due among other factors to their having a typical high degree of suggestiveness towards "Spring," (as their reasons for choices bring out,—regarding Spring clothing, etc.). Their broader range of variation seems (from the reasons given) to have its foundation upon a generally wider imagination, as regards the above Spring advertisements. Spring is in many respects a more decidedly outstanding season of the year for girls than for boys, and gives them more suggestions, thus making the degree of certainty in choice less.

Comparing our results in this test with the results of Series C., in which suggestiveness of a concrete object was involved we see that the sexes there agreed with each other to a much greater degree (0.79 as against 0.50 in Series S), and thus were on more common ground in their judgments. Also in Series C. the girls have a greater variability among themselves than the boys (girls ± 0.25 , boys ± 0.21), which seems to restate the fact of their having a greater general range of suggestiveness, and therefore, a lesser degree of certainty in placing the pictures by actual votes.

From the reasons given for choice it is clear that one and the same picture has been given the same rank by different observers, often for reasons having very little relation to each other.

Considering the factor of suggestiveness brought into an aesthetical judgment then, we may state, that a very complex process of judgment evolves which is based upon a variety of reasons.

SOME REASONS FOR CHOICE

Girls:

- (i) First appearance of his new straw hat.
- (j) Lambs and young birds most probable in Spring.
- (a) Young chickens just hatched.
- (d) *She* dressed in Spring clothes, but it is not warm enough yet for *his straw hat*.
- (g) Spring millinery.
- (f) Weddings more frequent in Spring.
- (e) Unsatisfied with winter hat,—he is too warm for it.
- I. (f) The huge raindrops I at once associate with April weather. Then, too, the Spring look upon the young man's face helps to carry out the idea of April showers.

- (j) The thing which suggests Spring most in this picture is the birds. The gamboling lamb is also suggestive of Spring. The hare one associates with Easter.
- (a) Chickens are for the most part suggestive of Spring. The rabbit and chick one connects with Spring.
- (c) The hare trying on a new hat and gazing in the mirror remind one of the usual custom of buying a new Spring hat.
- (d) The ladies' skirts blown by the brisk Spring winds.
- (c) The young chap is so glad that Spring is here that he climbs up into the newly leaved tree.
- (a) The young man is presenting someone with a straw hat; however, I usually associate a straw hat with Summer rather than Spring.
- (f) It is an old saying that in Spring young men lightly turn to thoughts of love; this sketch is, however, more suggestive.

Boys:

- II. Spring to me means first showers, then flowers, thus ("f" was most suggestive of rain, then spring lambs, in "j," came next, "c" with birds, flowers and boys came third, with golf came fourth, little children fifth. The others mostly suggested Spring styles.
- III. (c) The airiness.
- (d) The opening of feathers like awakening of Spring, with young people taking part.
- () Volubility of birds and light-heartedness.
- (b) Bright flowers.
- (a.f.h.) Excitement and joy of Spring. The first two, harmony.
- (i) Awakening of interest: blossoming out in new hat.
- (e) No reaction.

EXPERIMENT XVIII

Series AA. Judgment based on picturesque effect in *nine* scenes illustrating Dutch costumes of similar rich colorings in the foreground. As in the foregoing experiments the observer's attention was explicitly directed to one factor, here to the picturesque.

Average Rankings:

Labels of pictures	Girls	Boys
1	7	6
2	9	7
3	8	9
4	2	2
5	4	3
6	6	8
7	1	1
8	5	4
9	3	5

Individual correlations with their respective sex-group rankings:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ga
0.28	0.10	0.90	0.02	0.55	0.83	0.81	0.75	0.80	0.71	0.83	0.41
CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.42	0.77	0.53	0.63	0.70	0.87	0.43	0.67	0.63	0.33	0.60	0.40

Average Correlations Average Deviations

G.	0.59	± 25
B.	0.58	± 13

Correlation between girls and boys as one group=0.87.

Also here there is a very distinctly greater agreement between the sexes, than within them, one of the highest throughout.

The high uniformity of judgments, and the comparatively small average deviations, as well as the distinctly high correlation between the sexes as one group, indicate that there has been a very similar high uniform aesthetic reaction, pointing to a clear-cut impressionistic response.

From the reasons given for choice it may be inferred that the picture is generally reacted to as a whole, as some sort of an immediate, undefinable oneness, (which fact was also corroborated by the oral comments of the observers.) The observers were at a loss in *searching for reasons* after the choice was made and reflected longer before doing so. They all expressed spontaneous aversion and some positive inability to analyze their motives.

Words like "setting," "natural surroundings," "color effect," "atmosphere," etc., all suggest generic homogeneous reaction.

*Some typical reasons for choices:**Girls:*

- I. (a) Presence of windmills—extreme dress of woman and child—coloring.

- (g) Sheep—Naturalness of woman.
- (c) Quaintness of girls—bright coloring—background.
- (d) Water scene—dainty coloring—quaint figures.
- (h) Little town in distance—rather elaborate dress of woman.
- (e) Coloring—old age and youth.
- (b) Sameness of coloring—children's activities—boats.
- (f) Simplicity.
- II. (a) Few strokes, atmosphere.
- (b) Atmosphere and motion.
- (c) Atmosphere.
- (d) Motion.
- III. Judgment was based on color, scenery, characterization.

Boys:

- I. (a) 1 I believe the women are in a typical attitude—that is, of resting just before going to or leaving the market, the row boat in the background and the sail boat adds picturesqueness to the whole.
- (g) 2 The occupations of the people shown here in their environment is pleasing.
- (i) 3 The three figures in the foreground give one an impression of peace and industry—the manner in which they have squatted themselves on the sand is interesting.
- (h) 4 The bright colors in the aprons and skirts is pleasing.
- (f) 5 The plain rough clothing worn by the woman harmonize with the fishing ships in the background.
- (e) 6 The coloring here, although bright, is pleasing and the gradual dwindling of the line marking the water and sky gives me the impression of great distance.
- (c) 7 The prosperity of the country is evident. The sail boats as well as the green grass suggest picturesqueness.
- (a) 8 There is nothing very striking about this picture, except the odd costume.
- (b) 9 I do not like the tranquil, modest expression and attitude of the two young women.

- II. (a) 1 The most picturesque costume appeals to me first—local peculiarities of dress (first group.)
 (b) 2 Peculiarities of surroundings, quaint scenery, sheep raising (rural), fishing town—(second group.)
 (c) 3 Striking color presentation, unusual contrast of pleasing colors (third group.)

The reaction reminds one of the definition of William James of "Extensivity" as a primary basic element in sensation; the "vastness" or "bigness" of a primary visual impression, as given by him; the extensivity of sound volume as introduced by Watt. Is there a similar, primary, sensory, aesthetic reaction?

EXPERIMENT XIX

Series BB.—Judgment based on Preference, with a view toward permanent possession in observer's room.

Material.—Colored reproductions of famous paintings.

AVERAGE RANKINGS

Labels of pictures	Girls	Boys
1	2	4
2	4	9
3	8	7
4	9	2
5	6	6
6	1	1
7	7	10
8	10	8
9	3	3
10	5	5

Individual correlations with the average rankings of the sex-groups respectively:

fm	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ga
0.35	0.44	0.84	-0.35	0.60	-0.14	0.70	0.87	0.62	0.88	0.76	0.24
CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.50	0.58	0.48	0.50	0.19	0.09	0.45	0.59	0.19	0.82	0.88	0.15

Average Correlations of the Sexes as Separate Groups

Girls	0.48
Boys	0.45

Average Deviations

± 0.31
± 0.20

Correlation between girls and boys as one group=0.44.

The above results demonstrate that when the aesthetic judgment, as here, is narrowed down to sharp individual taste, and the stimulus pictures have very little in common (except that they are all famous paintings of men and women,)—as might have been expected,—the individual differences stand out distinctly.

The reasons given range from "the picture that can give me ideas, to help me be of better service in the world," to "color, light and shade; technique," with all varieties in between.

EXPERIMENTS XX—XXI

Judgment based on most effective portrayal of the feminine characteristic, Series CC, and, judgment based on most effective portrayal of the masculine characteristic, Series DD.

Material.—Ten representative portraits of each sex.

Average rankings by the girls and boys, as separate groups, respectively:

<i>Series CC</i>			<i>Series DD</i>		
Labels of Pictures	Girls	Boys	Labels of Pictures	Girls	Boys
1	1	5	1	2	3
2	8	7	2	9	10
3	3	6	3	7	6
4	2	4	4	10	9
5	9	9	5	3	2
6	7	8	6	1	1
7	5	1	7	4	4
8	10	10	8	8	8
9	6	3	9	5	5
10	4	2	10	6	7

Individual correlations with the average ranking of the *Girl* group:

	mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	ge
CC	0.27	0.43	0.65	0.59	0.50	0.70	0.25	0.71	0.79	0.64	0.73	0.49
DD	0.31	0.90	0.83	0.22	0.49	0.60	0.79	0.44	0.87	0.55	0.79	0.61

Individual correlations with the average ranking of the *Boy* group:

	CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
CC	0.61	0.36	0.18	0.48	0.61	0.98	0.73	0.79	0.01	0.28	0.82	0.58
DD	0.55	0.95	0.89	0.90	0.68	0.41	0.84	0.82	0.19	0.62	0.77	0.68

The average correlation of the girls as a group in Series CC = 0.56.

The average correlation of the girls as a group in Series DD = 0.62.

The average correlation of the boys as a group in Series CC=0.49.

The average correlation of the boys as a group in Series DD=0.69.

Series CC

Average correlations of the sex groups with their respective group rankings		
		A. D.
Girls	0.56	± 0.15
Boys	0.49	± 0.28
Correlation between the sexes as one group = 0.64		

Series DD

Average correlations of the sex groups with their respective group rankings		
		A. D.
	0.62	± 0.18
	0.69	± 0.17
Correlation between the sexes as one group = 0.96		

Again greater agreement between the sexes than within them,—in both series,—is observed. These two experiments bear out a greater agreement in regard to the masculine characteristic (Series DD) than in regard to the feminine characteristic (Series CC) within each sex group. And still more striking is this agreement when the sexes are taken together as one group and compared—this correlation being the highest of any experiment throughout all twenty-two experiments.

In Series CC there is a greater agreement on the part of the girl group. In Series DD there is somewhat greater agreement on the part of the boys. Also the difference in agreement on the same sex is approximately equal.

The boys are much more in agreement on their own sex, than the girls are on theirs.

From the reasons given the choice of the girls were based mainly on feminine charm and tenderness; and on motherliness by both boys and girls. Strength, determination and courage are the bases of judgment for the masculine characteristic by both sexes.

In regard to the agreement of the individual girl or boy with chance neighbors the following tables give a detailed account:

REYMERT

CORRELATIONS WITH CHANCE NEIGHBORS (Girls)

Series CC.

	mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
mf.....	0.43	0.43	0.43	0.04	0.03	-0.10	-0.10	0.08	0.05	0.47	0.01	-0.19
me.....	0.43	-0.01	-0.01	-0.05	-0.44	0.37	0.32	0.21	0.45	0.33	0.15	0.05
tk.....	0.04	-0.05	0.62	0.62	0.48	0.00	-0.14	0.41	0.30	0.53	0.30	0.05
ec.....	0.03	-0.44	0.62	0.47	0.47	0.41	0.08	0.53	0.45	0.36	0.19	0.09
ml.....	-0.10	0.37	0.00	0.44	0.35	0.35	0.15	0.41	0.32	0.22	0.62	0.31
ad.....	-0.10	0.32	-0.14	0.08	0.15	0.15	0.15	0.09	0.60	0.35	0.61	0.27
ar.....	0.08	0.21	0.41	0.53	0.32	0.41	0.09	0.90	0.41	-0.60	0.55	0.09
gl.....	0.05	0.45	0.30	0.45	0.32	0.60	0.41	0.90	0.90	0.40	0.43	0.25
mr.....	0.47	0.33	0.53	0.36	0.22	0.35	-0.60	0.40	0.28	0.28	0.58	0.33
bt.....	0.01	0.15	0.30	0.19	0.62	0.61	0.55	0.43	0.58	0.21	0.21	0.44
ch.....	-0.19	0.05	0.05	0.09	0.31	0.27	0.09	0.25	0.33	0.44	0.42	0.42
gs.....												
S.....	0.14	0.21	0.28	0.30	0.31	0.32	0.17	0.38	0.42	0.33	0.37	0.21
11												

Sum Correlations equals 3.44

CORRELATIONS WITH CHANCE NEIGHBORS (Boys)

Series CC.	CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
CC.....	0.12	0.12	0.68	0.70	-0.08	0.55	0.15	0.77	-0.36	-0.42	0.22	0.79
GJ.....	0.68	0.05	-0.05	0.20	0.19	0.28	0.41	0.35	0.53	-0.59	0.08	-0.19
HH.....	0.70	0.20	0.24	0.24	-0.16	0.15	-0.25	0.33	-0.25	-0.49	-0.24	0.55
LB.....	-0.08	0.19	0.24	-0.10	-0.10	0.47	0.15	0.66	-0.08	-0.42	0.18	0.48
CS.....	0.55	0.28	-0.16	0.47	0.63	0.63	0.76	0.09	0.41	-0.08	0.60	-0.06
NE.....	0.15	0.41	-0.25	0.15	0.76	0.67	0.67	0.71	-0.05	-0.31	0.79	0.59
LH.....	0.77	0.35	0.33	0.66	0.09	0.71	0.24	0.24	0.27	-0.41	0.50	0.71
WD.....	-0.36	0.53	-0.25	-0.08	0.41	-0.05	0.27	-0.39	-0.39	-0.08	-0.11	0.67
EO.....	-0.42	-0.59	-0.49	-0.42	-0.08	-0.31	-0.08	-0.41	-0.08	0.18	0.18	-0.16
DN.....	0.22	0.08	-0.24	0.18	0.60	0.79	0.78	0.50	-0.11	-0.16	0.38	0.38
RG.....	0.79	-0.19	0.55	0.48	-0.06	0.59	0.09	0.71	-0.67	-0.16	0.34	0.33
EK.....	0.36	0.20	0.05	0.23	0.20	0.41	0.32	0.32	0.18	-0.28	0.34	0.33
S.....	0.36	0.20	0.05	0.23	0.20	0.41	0.32	0.32	0.18	-0.28	0.34	0.33
11	0.36	0.20	0.05	0.23	0.20	0.41	0.32	0.32	0.18	-0.28	0.34	0.33

Sum Correlations equals 2.30

EXPERIMENT XXII

Series EE.—Judgment based on degree of impressiveness of ten religious pictures.

Average rankings by the sex groups, respectively:

Labels of pictures	Girls	Boys
1	1	2
2	9	10
3	8	5
4	6	9
5	4	1
6	5	4
7	10	6
8	2	3
9	7	8
10	3	7

Individual correlations with the average ranking of the *Girl* group:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
0.27	0.76	0.38	0.54	0.67	0.37	0.28	0.88	0.41	0.62	0.15	0.22

Individual correlations with the average ranking of the *Boy* group:

GC	CJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.62	0.50	0.32	0.36	0.61	0.60	0.42	0.54	0.54	0.78	0.41	0.10

Average correlations of girls
and boys as separate groups

Girls	0.46
Boys	0.41

Average Deviations

± 0.19
± 0.23

Correlation between girls and boys as one group = 0.61.

As might have been expected, there is a great general diversity in choice here. Looking over the reasons given for the main choices by the girls:

"Christ's Humility for the Redemption of the World."

"The helpless are protected." "Motherhood."

"Biblical significances and symbolic Characters."

"Death, suffering, sorrow," etc.

We see how outspokenly they have been determined by biblical religion.

The reasons by the boys:

"Impressive surroundings conveying beautiful thoughts."

"Biblical significance." (The symbolism.)

"Death."

"Portrayal of suffering."

"Sympathy and fear," etc.

Biblical symbolism seems to play the greater part in all their judgments, and varies with the individual temperaments.

To get a somewhat *rough* measure of the degree to which the different girls and boys vary around the group standard judgment throughout the twenty-two experiments we may take the *average* of each individual's correlations with his sex-group in each experiment and compare this with the *average* of all twenty-two standard judgments of the sex-groups, as below:

AVERAGE CORRELATIONS WITH THE GROUP

(throughout all twenty-two experiments)

FOR THE INDIVIDUAL GIRLS AND BOYS

Girls:

mf	me	tk	ec	ml	ad	ar	gl	mr	bt	ch	gs
0.30	0.14	0.43	0.75	0.65	0.55	0.48	0.37	0.71	0.68	0.38	0.49
0.77	0.43	0.76	0.71	0.78	0.78	0.84	0.59	0.82	0.71	0.76	0.77
0.75	0.72	0.70	0.08	0.75	0.49	0.88	0.59	0.81	-0.31	0.77	0.88
0.53	0.85	0.85	0.25	0.53	-0.27	0.61	0.04	0.26	-0.04	0.02	0.54
0.88	0.84	0.52	0.73	0.76	0.28	0.87	-0.32	0.32	0.95	0.33	0.41
0.59	0.79	0.33	0.75	-0.05	0.84	0.77	-0.44	0.16	0.81	0.43	0.45
0.38	0.84	0.10	0.77	0.30	0.68	-0.22	-0.27	0.37	0.96	0.65	0.98
0.09	0.50	0.76	0.59	0.64	-0.45	0.65	0.47	0.84	0.65	0.78	0.25
0.71	0.83	0.84	-0.03	0.49	0.10	0.05	-0.07	0.07	0.81	0.09	0.13
-0.05	0.87	0.15	0.15	0.94	0.05	0.03	-0.05	0.05	0.92	0.12	-0.03
0.66	0.99	-0.20	0.98	-0.59	0.66	0.89	0.66	0.95	0.60	0.66	0.76
0.59	0.64	0.88	0.25	0.53	0.68	0.87	0.76	0.41	0.50	0.66	0.26
0.55	0.66	0.27	0.28	0.40	0.68	0.48	0.81	0.82	0.81	0.61	0.73
0.75	0.90	0.43	0.58	0.92	0.49	0.81	0.64	0.67	0.13	0.79	0.31
0.77	0.50	0.77	0.82	0.71	0.75	0.76	0.67	0.70	0.78	0.22	0.77
0.98	0.93	0.05	0.54	0.82	0.64	0.90	0.83	0.76	0.83	0.33	0.75
0.89	0.36	0.64	0.79	0.79	0.27	0.59	0.20	0.24	-0.09	0.67	0.90
0.28	0.10	0.90	0.02	0.55	0.83	0.81	0.75	0.80	0.71	0.83	0.45
0.35	0.44	0.84	-0.35	0.60	-0.14	0.70	0.87	0.62	0.88	0.76	0.24
0.27	0.43	0.65	0.59	0.50	0.70	0.25	0.71	0.79	0.64	0.73	0.49
0.31	0.90	0.83	0.22	0.49	0.60	0.79	0.44	0.87	0.55	0.79	0.61
0.27	0.76	0.38	0.54	0.67	0.37	0.28	0.88	0.41	0.62	0.15	0.22
S. 11.62	14.42	11.88	10.01	12.18	9.58	13.09	9.13	12.45	13.10	11.53	11.36
C. 0.53	0.66	0.54	0.46	0.55	0.44	0.60	0.42	0.57	0.60	0.52	0.52
A.D. 0.23	0.21	0.27	0.29	0.22	0.30	0.26	0.35	0.25	0.26	0.24	0.23

Grand average correlation of the individual girls=0.53

Grand average deviation of the individual girls= ± 0.26

From the table we again see the general high correlation of each individual subject, both girls and boys, with their respective group judgment (the span of the girls being 0.42 to 0.66,—the span of the boys from 0.31 to 0.65.) The individual variability around the group judgment is on the whole the same for both sexes although the span for the girls is between 0.21 to 0.35, as against that of the boys, 0.16 to 0.35.

The average correlations of the group judgments of the sexes as separate groups are both high. However, the standard girl group correlation is higher than that of the boy group, which all goes to demonstrate that the girls agree better among themselves than do the boys throughout the twenty-two experiments in preference judgments of the kind here presented.

Boys:

CC	GJ	HH	LB	CS	NE	LH	WD	EO	DN	RG	EK
0.59	0.25	0.90	0.45	0.70	0.88	0.62	-0.30	0.90	0.24	-0.13	0.40
0.89	0.55	0.52	0.76	0.89	0.13	0.98	0.85	0.44	-0.19	0.08	0.21
0.71	0.72	0.72	0.82	0.62	0.41	0.62	0.38	0.41	0.62	-0.36	0.64
0.85	0.16	0.79	0.50	0.33	0.37	-0.02	0.28	0.50	0.14	0.38	-0.32
0.71	0.43	0.49	0.62	0.90	-0.07	0.75	0.35	0.42	0.66	-0.36	0.48
0.62	0.42	0.76	0.81	0.53	0.81	0.58	0.49	0.87	0.13	0.13	0.47
0.54	0.37	0.90	0.48	0.60	0.71	0.84	0.26	0.20	0.82	0.53	0.60
0.89	0.66	0.58	0.38	0.94	0.37	0.01	0.50	0.52	0.72	-0.37	0.18
0.82	0.66	0.62	0.10	0.41	-0.55	0.85	0.72	0.67	-0.19	0.19	0.36
0.84	0.36	0.82	0.78	0.28	0.81	0.84	0.61	0.94	0.07	0.37	0.55
0.70	0.81	0.28	0.90	0.89	0.77	0.14	0.78	0.88	0.70	-0.41	0.71
0.89	0.53	0.67	0.13	0.78	0.20	0.45	0.21	0.40	0.66	0.58	0.33
0.62	0.54	0.07	0.05	0.54	0.04	0.27	0.47	0.50	0.59	0.60	-0.19
0.24	0.22	-0.01	0.79	-0.03	0.52	0.68	0.47	0.13	0.04	0.73	-0.35
0.83	0.59	0.59	0.78	0.88	0.41	0.40	0.82	0.38	0.14	0.75	0.47
0.82	-0.07	0.26	0.98	0.87	0.95	0.92	0.89	0.33	0.48	0.09	0.64
0.07	0.52	0.50	0.03	0.27	0.85	0.49	0.28	-0.52	0.48	0.38	0.48
0.42	0.77	0.53	0.63	0.70	0.87	0.43	0.67	0.63	0.33	0.60	0.40
0.50	0.58	0.48	0.50	0.19	0.09	0.45	0.59	0.19	0.82	0.88	0.15
0.61	0.36	0.18	0.48	0.61	0.98	0.73	0.79	-0.01	-0.28	0.82	0.58
0.55	0.95	0.89	0.90	0.68	0.41	0.84	0.82	0.19	0.62	0.77	0.68
0.62	0.50	0.32	-0.36	0.61	0.60	0.42	0.54	0.54	0.78	0.41	-0.10
S. 14.33	10.16	11.86	11.51	13.19	10.56	11.21	11.47	9.51	8.38	6.76	7.37
C. 0.65	0.46	0.54	0.52	0.60	0.48	0.51	0.52	0.52	0.38	0.31	0.34
A.D. 0.16	0.26	0.21	0.27	0.21	0.32	0.24	0.21	0.21	0.31	0.35	0.25

Grand average correlation of the individual boys=0.49

Grand average deviation of the individual boys= ± 0.25

Grand average correlation between girls and boys as one group=0.68—with an average deviation of= ± 0.14 .

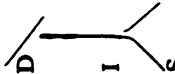
See also Tables, pages 133 and 134

The deviation from the standard sex-group average correlation is practically the same for both sexes as separate groups, but only slightly higher than one-half of this when the sexes are taken together as one group. The latter plus the standard group average correlation between the sexes as one group, 0.68, shows a very high degree of agreement between the girls and boys, the correlation of the mixed-group being throughout the twenty-two experiments a conspicuous substantiation of the fact of greater agreement between the sexes as one group than within them as separate groups.

TABLE SHOWING THE CORRELATIONS IN ALL TWENTY-TWO EXPERIMENTS AND THE CORRELATIONS OF RELATED GROUPS OF SERIES

Experiment Number	Series	Materials	For Girls as a group (g) For Boys as a group (B) Between Girls and Boys as one group (Gr.)			Average correlation between girls and boys as one group for the Charcoal Drawings Series=0.68
			g	B	Gr.	
1	A) B) C)	Simple pieces of pottery Pattern designs Different swirls	0.49	0.46	0.49	0.77
2			0.73	0.51	0.77	
3			0.59	0.53	0.79	
4	I) II) III) IV) V) VI) VII)	7 representative types of non-geometric figures, consisting of varying relations between two curved lines and two straight lines	0.35	0.33	0.84	0.71
5			0.55	0.45	0.65	
6			0.45	0.55	0.73	
7			0.46	0.57	0.88	
8			0.48	0.45	0.59	
9			0.34	0.39	0.33	
10			0.26	0.55	0.70	
11			0.59	0.60	0.94	

For all Proportion Tests=

Experiment Number	Series	Materials	g	B	Gr	
12	A & B	Simple preference Color Design	0.59	0.49	0.43	For the Magazine Cover group of Series=0.59
13	A & B (Color)		0.60	0.34	0.78	
14	A & B (Design)		0.62	0.29	0.56	
15	 Ceramic Art Designs	Non-geometric (Judged for design)	0.69	0.59	0.77	For Ceramic Art De- signs=0.74
16		Geometric (Judged for in- tricacy of design)	0.70	0.60	0.70	
17		Illustrated Spring adver- tisements	0.52	0.32	0.50	

Arbor and Bird
magazine
cover illus-
trations; colored

<i>Experiment Number</i>	<i>Series</i>	<i>Materials</i>	<i>I</i>	<i>B</i>	<i>Gr</i>
18	AA	Dutch Costume	0.59	0.58	0.87
19	BB	—Paintings	0.48	0.45	0.44
20	CC	Ten religious pictures	0.46	0.41	0.61
21	DD	Ten portraits of women	0.56	0.49	0.64
22	EE	Ten portraits of men	0.62	0.69	0.96
		Sum Average	11.72	10.64	14.97
		Average Correlation	0.53	0.48	0.68
		Sum average of Deviations	2.08	1.90	3.13
		A. D. for all series	0.09	0.09	0.14

For the Post Card Reproductions—0.66

Post Card
reproductions
(Colored)

From the above tabular arrangement we can readily gain a direct idea of the degree of correlation that was found between the various experiments that suggest a grouping with respect to relation and similarity in values tested or values that suggest types. The span is from 0.59 to 0.74 (neglecting experiments S and AA which show the extremes of 0.50 and 0.87.)

The following tabular arrangement of the correlations of the (1) girls, (2) boys, and (3) the sexes as one group, in gradation, according to degree of correlation, gives an oversight over the amount of general agreement there was among the three groups in their rating of the various aesthetic values experimented upon.

We see that for the three groups there is a span from 0.26 to 0.73 for the girl group, from 0.29 to 0.69 for the boy group, and one from 0.33 to 0.96 for the sexes as one group.

The first two have practically the same span, the girls a slightly larger, and the third group has the largest span.

GRADED CORRELATIONS

Girls	Boys	Gr.
Series	Series	Series
B.....0.73	DD.....0.69	DD.....0.96
I.....0.70	VII.....0.60	VII.....0.94
D.....0.69	I.....0.60	IV.....0.88
DD.....0.62	D.....0.59	AA.....0.87
A & B		
(Design).....0.62	AA.....0.58	1.....0.84
A & B	IV.....0.57	C.....0.79
(Color)		
AA.....0.59	III.....0.55	A & B0.78
A & B.....0.59	VI.....0.55	(Color)
VII.....0.59	C.....0.53	D.....0.77
C.....0.59	B.....0.51	B.....0.77
CC.....0.56	A & B.....0.49	III.....0.73
II.....0.55	CC.....0.49	VI.....0.70
S.....0.52	A.....0.46	I.....0.70
A.....0.49	II.....0.45	II.....0.65
BB.....0.48	IV(L).....0.45	CC.....0.64
IV(L).....0.48	BB.....0.45	EE.....0.61
EE.....0.46	EE.....0.41	(IV(L).....0.59
		A & B
IV.....0.46	V.....0.39	(Design).....0.56
III.....0.45	A & B.....0.34	S.....0.50
	(Color)	A.....0.49
1.....0.35	1.....0.33	BB.....0.44
V.....0.34	S.....0.32	A & B.....0.43
VI.....0.26	A & B.....	V.....0.33
	(Design).....0.29	

SUMMARY DISCUSSION

Series DD, involving the judgment of the masculine characteristic, shows the highest correlation, 0.96, between the sexes as one group, and is also the highest of all correlations. This may be due to greater popular agreement on the element suggested in this experiment.

Series VII. IV and 1, involved a very simple and clear proportion relation between two curved and two straight lines, and were therefore relatively easy to rank. Their correlations are 0.94, 0.88 and 0.84, respectively.

Series AA, involving a judgment of the picturesque effect, shows the high correlation, 0.87, between the sexes as one group. This may be due to the fact that this series, depicting Dutch costumes in a scenic setting, is popular, simple, rich in color effects, attractive, direct in its appeal, and so comparable that great difficulties in choosing were expressed by the subjects.

Series C, though not admitting of as ready judgment as the last named series, correlates very high, 0.79. The swirls in charcoal appeal to the imagination in a concrete, simple, direct manner.

Series A' and B (Color), the Arbor and Bird Day magazine covers, correlates similarly high, 0.78. The covers as their name would suggest were of general human appeal, familiar as well as attractive in subject, and illustrative of life out-of-doors and of the outdoor world. That the element of color played an outstanding part is obvious from the correlation of this series when judged on the basis of color.

Series D and Series B, the former involving judgment on beautifully colored non-geometric ceramic art designs, and the latter involving judgment on charcoal pattern designs, correlated high and equal, 0.77; this in spite of the fact of their great difference. The general attractiveness of the designs as well as their direct concrete impressionistic appeal and decorative suggestiveness may account for the results.

Series III and VI correlate comparatively high, 0.73 and 0.70, in the proportion tests, perhaps for similar reasons as those above.

Series I, in which the judgment was based on intricacy of design, correlates high, 0.70, especially when it is considered that a more complex element is involved. Here the designs were also beautifully colored ceramic art designs but geometric and involving great intricacy of design.

Series II, and IV (L) of the proportion tests have a medium

high correlation, 0.65 and 0.59 respectively, due perhaps to being somewhat less direct in type and having a greater chance for confusion.

Series CC, involving judgment of the feminine characteristic, shows only a moderately high correlation, 0.64. This may be due to a greater diversity of individual taste and ideals in regard to the feminine than in regard to the masculine characteristic, as borne out by the reasons assigned for choices.

Series EE, the judgment being based on the impressiveness of a set of reproductions of religious paintings, correlates medium, 0.61, contrary to expectations from the comments of the observers regarding their difficulty of choice. The biblically significant entered largely in this judgment.

Series A and B (Design) correlates only medium, 0.56, when judged on the basis of design, showing that design has a less general appeal than color—a racial development element as well as one of individual development.

Series S, which has its judgment based upon the suggestiveness of the idea of Spring in a set of Spring advertisements, involved a multiplicity of suggestions—the same picture being frequently chosen for altogether different reasons. Hence, as may be expected, the correlation is *comparatively* low, 0.50.

Series A and B, when a general simple preference judgment is passed upon it, correlates very much lower, 0.43, than when the judgment is based upon design,—showing clearly the presence of conflicting factors.

Series A, involved a judgment of technique of a set of charcoal drawings of two simple pieces of pottery drawn from various yet only slightly different angles. The low correlation, 0.49, was most likely due to the great homogeneity of this series.

Series BB, involving a judgment in regard to choice as a permanent possession in the room of the observer, had generically nothing in common, save that they all were colored reproductions of famous paintings. The correlation, consequently is low, 0.44.

Series V, which required the finest discrimination of the proportion tests, the increments of increase in width being of the smallest degree of all these tests,—(the figure is the same as in Series VI),—correlated only 0.33.

Series C, which is suggestive of a concrete object, and Series S, which is suggestive of an idea (Spring) bear an interesting relation to each other, $C=0.79$, $S=0.50$, showing the part the entrance of the intellectual element plays.

Series D and Series A and B (Design), which involved

design, pure and simple, and design in illustration, respectively, bear a somewhat similar relation to each other, $D = 0.77$, A and $B = 0.56$.

The span between the sexes as one group is greater than the span for either sex group separately:

Group = 0.33 to 0.96

Girls = 0.26 to 0.73

Boys = 0.29 to 0.69

The variability in correlations, as well as the general agreement on some series, is readily seen in the above tabulation.

A rough rating of the subjects from the *personal impression* gained by the experimenter throughout the twenty-two experiments on the degree of *conscientiousness*, *general intelligence* and *confidence of judgment*, is inserted below, for possible purposes of comparison.

	CONSCIENTIOUSNESS		GENERAL INTELLIGENCE		CONFIDENCE	
	Girls	Boys	Girls	Boys	Girls	Boys
Above Average	ch ad ar tk	RG CS CC HH WD	cs me ad gl	CS WD RG CC HH	ch gl ad ar me	CS RG HH NE DN CC GJ WD
Average	bt ec me gl mr	GJ DN NE EO LH LB	ar tk ec mr mf bt	DN NE GJ EK	tk bt mf ec	
Below Average	ml gs	EK	ml gs	LH EO LB	mr ml gs	LH LB EO EK

A SHIFT OF EMPHASIS NEEDED IN PERSONNEL RESEARCH

By HARRY D. KIRSON, Professor of Psychology, Indiana University

The term personnel research means to most people, to whom it means anything, the administration of tests—intelligence, “special ability,” and trade tests—to persons applying for positions. As a matter of fact this is a much too narrow view. The analysis of the individual (for which tests are the generally accepted though perhaps not the ultimate implements) is only half, and the latter half of the business. A task of equal importance is the analysis of the occupation to which the individual is to be adjusted. This has been almost entirely neglected, however. The cult of tests for analyzing the individual, with its formalized ritual, has so dazzled its devotees as to blind them to the possibility of another method of approach to personnel problems. The unfortunate consequences are not merely the neglect of the indispensable analysis of occupations but also a stagnation in the development of tests.

For it must be admitted that despite the great attention paid to tests, they are contributing but little to industry, in terms of either bulk or quality of service. There is much talk about them, but it is in terms of what might possibly be accomplished if they were developed along the line of a particular theory of vocational guidance. The few cases in which they have been practically used are conspicuous because of their rarity; and have not generally been followed up as a regular matter of industrial routine. One who views the matter dispassionately, then, must admit that in spite of the attention tests have received (in talk) they have not made any appreciable progress in industry during the past five years. Despite the rose-colored dreams that are painted of an industrial system wherein each individual shall be placed on the road to success by means of a dramatic series of psychological tests, the application of tests to occupational situations consists in following a stereotyped prescription that brings no cure and no change. In the psychologist's own jargon it may be said that psychological tests in industry are on a distinct plateau.

In order to escape from this plateau the industrial psychologist may well follow the prescription given by his colleagues in the psychology of learning, namely, change his method. Instead of seeking to solve the problems of personnel work by testing the individual with artificial, extraneous tests, let him approach the operations of the occupation as the worker actually has to perform them. By doing so he will not only fulfill the demands of an adequate program of personnel research, but he will also probably unearth some method that will show the way off the plateau upon which tests now languish.

This idea of applying psychological technique to the analysis of occupations is not new in theory. The public interested in vocational guidance have urged it under the name, vocational analysis, meaning the analysis of the professions and trades; or under the name, occupational analysis, meaning the analysis of specialties within the trades. Within recent times employment managers have been urging it under the name Job Analysis, meaning the analysis of a particular operation on a particular piece of material or a particular machine. Almost all who have faced the problems of personnel administration have recognized the necessity of this analysis, and particularly its psychological nature, but very few have actually attempted it, and they with only scattered efforts. As usually conceived, the psychological analysis of an occupation¹ consists in describing the occupation by means of rough observation on the part of a psychologist in consultation with employment manager, foreman, expert workman and others. The psychologist may even perform the operation more or less haltingly himself in order to secure "local color." After some such rule of thumb observation the occupation is split up into a series of "unit operations." These are then embodied, for purposes of selecting fit workers, into a statement called the Job Specification, for which outlines of various kinds exist.² An analysis according to such a pattern may be of value for some occupations and some circumstances, but it is surely inadequate

¹ It is of course understood that a complete analysis of an occupation requires description in terms of the physical, physiological, social and economic phases as well as of the psychological.

² See, for example, H. C. Link, *Employment Psychology*, New York. Macmillan, 1919, p. 261.

Ordway Tead and Henry C. Metcalfe, *Personnel Administration*, New York, McGraw & Hill, 1920, p. 264-5.

Franklyn Meine, *Job Specifications*, Bull. No. 45, Fed. Bd. for Voc. Education, Washington, D. C., Nov. 19, 1919.

as a sole instrument of scientific personnel research in the broad sense. It should be amended in accordance with three principles of scientific method:

1. Analyze the occupation in terms of itself.

As at present conducted, most attempts at occupational analysis violate this in one of three ways.

- a. According to one method greatly favored, the psychologist gazes at the occupation for a short time and then concludes: "This occupation requires concentration of attention, verbal memory, reasoning about new situations, etc. Here are tests a, b, and c, which measure these several activities (sic!). I will administer these tests to workers in the occupation and choose those tests wherein scores show the highest correlation with standings in occupational success." The "tested" activities *constitute* the activities of the occupation, in the eyes of the analyst; and the procedure is regarded as a piece of occupational analysis. Upon second thought the fallacy of this procedure is grossly apparent. It really delineates the occupation only in terms of extraneous tests, not in terms of the work itself. And it can be justified only on the obsessing hypothesis that tests with their wabby indices of correlation are the most reliable tools of analysis.

- b. Another practise much in vogue is to make a list of psychical traits supposedly required in an occupation—industriousness, affability, neatness, patience, assertiveness—and call them the analyzed components of the occupation. The chief objection to the listing of such traits is that they are abstract, characterizing a number of occupations and yet no occupation in particular. Again they are general; every occupation requires them in some degree. Finally, it is not yet established that a trait like neatness, for example, will express itself equally in all lines of activity. The greater likelihood is that each occupation involving it will require it in a specific form.

A good example of an attempt to overcome this last objection is furnished by an insurance company that desired to analyze out the psychical characteristics brought into play in the best quality of insurance salesmanship. Having discovered by trial the futility of using such general terms as those just mentioned, the company has attempted to formulate the required traits in terms of what the best salesman does in selling insurance. Thus, instead of scheduling a mythical, abstract, general trait like "dominatingness," the analyst asks

for a rating in this manner: "Does he (the best salesman) dominate the interview most of the time; half of the time; almost none of the time?" This method is not perfect; it involves a danger to be mentioned presently; it is only tentative; but it has the virtue of avoiding general terms that mean nothing; and it shows a praiseworthy attempt to analyze the occupation in terms of itself.

c. A third questionable practise among the current attempts at occupational analysis is that of seeking the components of an occupation in the opinions of expert workers. One of the most thoroughgoing applications of this technique is that of Martha Ulrich,³ who has devised a schema consisting of a number of traits, which she presents to professional men—doctors, lawyers, etc.—with the request that they rate each trait according to the degree to which they think it essential in their occupation. Thus:

1 desirable	—1 undesirable
2 very important	—2 disadvantageous
3 absolutely indispensable	—3 very disadvantageous
0 unimportant	

More than 100 traits are scheduled, of the sort condemned in (b) above, such as: visual, auditory or motor type of imagery; ability to make new associations easily; ability to think in abstract terms; a strongly developed emotional life, etc. The judgments of all are then summed up to give a numerical weight to each trait as a component of each vocation.

This technique in its various guises constitutes a most precarious tool of vocational analysis. The dangers of subjective data that lurk in it emphasize anew the need for the kind of analysis urged in this paper, namely the analysis of an occupation in terms of itself instead of in terms of psychological tests, abstract non-existent general "traits," or opinions.

2. The second principle is, Express the requirements of the occupation in quantitative terms. For example, in describing the occupation of proofreading, one should, among other things, measure the output of a standard group of workers, and upon this basis, state the median requirement to be a speed of 900 lines per hour (of a certain kind of material) with no more than one error overlooked every 35 lines. Or, to change the illustration, one might state the speed required

³ *Die psychologische Analyse der hohen Berufe als Grundlage einer kuenftigen Berufsberatung*. Zeitschrift fuer Angewandte Psychologie, 13, 1918, pp. 1-36.

in linotyping at the upper ten percentile, to be 8,300 lines per hour with no more than one error in every 38 lines. Or one might describe typewriting of grade A as an output of 100 letters per day, B, 80 and C, 60.⁴

3. A third requirement is that the analysis should be made minutely. Perhaps the rough, non-quantitative kind now practised is adequate as a basis for selecting workers for certain gross unskilled occupations. But when one takes the pains to measure the reaction time of an applicant in fractions of a minute as one does in tests, one should have measured the object to which the worker must react (the occupation) with equal minuteness.

There are still other considerations that demand that occupational analysis should be made minutely. Analysis is necessary for purposes other than merely selection of workers. It may be used for discovering wasteful and fatiguing operations, as when Taylor found that to analyze such a gross job as shoveling required measures involving ounces; and Gilbreth found that to analyze the folding of handkerchiefs required chronocyclegraphs showing movements lasting but a fraction of a second. Again, analysis may be made for the purpose of discovering what factors are involved in learning an occupation, as in the analysis of telegraphing by Bryan and Harter and that of typewriting by Book. A more recent attempt by the writer in the field of proofreading illustrates the same requirement. The particular aim of the investigation was to secure a standard of eye-movement. The established technique for the measurement of eye-movements was employed, which permitted their reproduction in terms of the number and length of pauses. Table I shows these averages in fiftieths of a second for good and poor proofreaders reading the same twelve lines.

TABLE I

	Average number of pauses per line	M. V.	Average length of pauses	M. V.
Good	7.4	.48	17.1	1.66
Poor	11.0	2.00	19.3	3.64

"From these results we may draw the following conclusions: (1) a good reader makes on the average fewer pauses than a poor one—7 *vs.* 11. (2) A good reader makes on the

⁴ Here again, the reader is reminded that the same kind of quantitative statements should be made on behalf of the non-psychological characteristics of an occupation.

average shorter pauses than a poor one—17 *vs.* 19. (3) A still more important matter is the demonstration of differences in regularity. The good proofreader reads very regularly. His eye-movements are so regular as to be almost rhythmic, with respect to both length of line and number of pauses. The poor proofreader, however, reads irregularly. These differences may be quantitatively stated in terms of the mean variations. The small mean variations in the record of the good reader indicate that he paused practically the same number of times in each line—on the average no more than eight ($7.4 \pm .48$) times, and no less than seven ($7.4 - .48$) times. The poor reader, however, paused in one line as many as thirteen (11.0 ± 2.0) times and in another line only nine ($11.0 - 2.0$) times. In length, the pauses varied between 18.8 and 15.4 in the case of the former, and between 22.9 and 13.7 in the case of the latter.”⁵

RÉSUMÉ

Such are the three principles that should be followed out in any adequate program of occupational analysis; and such are the methods that should be employed. This rigid régime is admittedly arduous. It requires that the psychologist who would engage in personnel research must leave the shelter of his academic laboratory and enter the arena of occupations. He need not necessarily don a mechanic's cap or a mason's apron, but he must seek his data in the activities of the worker at work. He need not necessarily invent a battery of new instruments for the measurement of these work-activities. He may secure many of the necessary measures from the records of well-organized industrial and commercial establishments. There are vast quantities of such measures available to one who will gain the confidence of live business executives and show them by results that personnel psychologizing as described above will benefit not only personnel work in general but also their establishment in particular.

It may be objected that this program of minutely measuring occupations in terms of themselves is going to take a long time, and that the work of vocational guidance cannot wait, hence must proceed with the dubious makeshifts (particularly tests) described above. The answer is that to search for truth always takes time but is more economical in the long run. Or still more tartly, that tests are not leading to

⁵ H. D. Kitson, *Scientific Method in Job Analysis*. *Jour. of Pol. Econ.* 29, June, 1921, p. 512.

speedy victory, and judging from the length of the plateau they are now on, cannot progress much farther in their present form. Hence, the most economical thing to do is to *lessen the emphasis upon tests, at least temporarily, and place more emphasis upon the scientific analysis of the occupation until it has reached a stage of refinement equal to that of the tests.* Then, perhaps, the way will be clear to improve the tests.

Again the objection may be raised that even after an occupation has been pretty well analyzed in terms of itself, the way will not have been opened for testing the ability of a novice aspiring to the occupation, for of course one could not be expected to perform the skilled operations required in the occupation before learning them. It is freely acknowledged that the reforms here advocated do not seem to offer a ready solution to the problem of selecting a candidate for an occupation by tests. Still it is possible that after a truthful analysis of the occupation is made, the way may be open for the development of the needed technique of vocational selection and guidance. One who would be bold enough to defy conventions might suggest further that perhaps it will turn out that tests are not the best tools to use after all. It is conceivable that a better way may evolve out of a well-directed search for light.

All apart from its bearing upon tests, however, the scientific analysis of occupations here advocated is bound to result in a number of benefits to personnel work conceived in its broader sense: 1. It will reveal errors and wastes that may be present in an operation, thus reducing the worker's fatigue and bringing profit to both worker and employer. 2. It will furnish basis for establishing standards of performance that may be demanded of employees. 3. It will assist in scientifically setting a wage consistent with the amount of work done. 4. In connection with jobs involving the use of machinery it may give information regarding more effective ways of adjusting, operating and maintaining machines. 5. The cost department may use accurate measures of the units of an occupation in estimating more closely the cost of new jobs about to be contracted for. 6. Analysis may conceivably be used in arranging harmonious grades and routes of promotion, showing the points of overlapping in the duties and qualifications of various jobs, and thus indicating at what stages workers traveling various routes may be interchanged. 7. Scientific analysis is an indispensable preliminary to the invention of a trade test. 8. Finally, it is a prerequisite to adequate occupational training.

In short, a scientifically defensible theory and administration of occupational analysis will result in many benefits not specifically foreseen; as happened in England when the Industrial Fatigue Research Board, organized during the war in order to "promote better knowledge of the relations of hours of labor and of other conditions of employment including methods of work to functions of the human body," discovered facts and developed techniques that have since been found directly applicable to the solution of problems other than those at first proposed.

The secret is that under the plan here advocated, the personnel psychologist will be working not with spurious extraneous materials, but with the actual stuff of occupations. And when he sheds obsessing theories and seeks for facts in the realm of reality he is bound to discover truth even if not in the form expected.

THE USE OF THE PERSONAL HISTORY BLANK AS A SALESMANSHIP TEST

By DOROTHY B. GOLDSMITH, The Guardian Life Insurance Company
of America

This study was made to determine whether the items of a personal history blank could be used to predict the success or failure of a salesman. That is, whether a weighted, quantitative score could be given to the answers to questions upon a history blank, in order that it might be used as an elimination test. To accomplish this, the scoring of the blank had to perform two functions,—(1) it must eliminate failures, and (2) it must *not* eliminate successes. The chief purpose was, therefore, the establishment of a critical score, a score below which would fall the failures and above which would lie the successes.

The attempt embodied in this study was successful; a lower critical score was established at a point which for the 502 blanks studied, eliminated 54% of the failures, but left 84% of the successes with the company. For this group the establishment of an upper critical score was unnecessary.

In this experiment the personal history blanks of 502 salesmen of The Guardian Life Insurance Company of America were studied. For the life insurance field in particular, though it would hold good in some measure in the case of any large organization, there is a peculiar advantage in adapting the personal history blank to this purpose. In this business, branch offices are maintained all over the country, and it is an extremely difficult task to persuade the branch managers to use tests, and a well nigh impossible one to have the tests given under standardized conditions. The personal history blank, however, has for some time been a routine part of the application for a sales license. As its name implies, it contains information concerning the applicant's personal history and past experience. Almost every firm uses a blank of some description and its completion presents no novel task. The manager has become accustomed to having this blank completed by his new men, and, in turn, to sending it on to the Home Office. In the past, this blank had been required for every agent of the company under discussion. There had

been no attempt, however, to give it any scientific treatment; the personal history blank had served rather, in the main, merely to strengthen or weaken the general impression made by the applicant. Certain items whose presence or absence were deemed significant were checked, and the remainder of the blank was taken as a whole to determine by its general tone whether an applicant should be licensed or dropped as a possibility.

To secure the proper scoring for the blank, the results of a study made about a year ago on 34 salesmen were taken as a tentative guide.¹ Fifty history blanks representing three classes of agents whose production records were known were selected. These classes were (1) those who were failures, (2) those who were borderline cases up to moderately successful, and (3) those who were successful. The criterion of success was the amount of insurance paid for during the first year after the man was licensed. (We had already found out that the first year's production was a very good index of future production.) It would be well here to mention that the only factor which determined the selection of the 502 blanks was whether or not the agent's first year's production record was available. The blanks for these three classes of men were then studied and the essential or significant items which varied with the success of the agent were selected. Each blank was then graded according to the tentative scoring, and the total thus obtained was checked against the man's production record. In accordance with these results, the scoring was revised, and the same process repeated with 25 blanks selected at random. We found that those whose scores on the blank were low were also low in production, and that with increasing production the scores tended to be higher. The correlation, of course, was not a perfect one, but we were interested rather in establishing the lower critical score than in correlations. Inasmuch as the personal history blanks selected at random were well distributed into our three classes by the scoring which we had determined upon, we selected this scoring as our final method. The weighted values thus assigned to the significant items were as follows:

<i>Age</i>		28-29	+2
18-20	-2	30-40	+3
21-22	-1	41-50	+1
23-24	0	51-60	0
25-27	+1	Over 60	-1

¹ Report A—1921, Scoring of Personal History Blanks. Bureau of Personnel Research, Carnegie Institute of Technology.

Marital Status

Married +1
Single -1

Occupation

Social +1
Unsocial -1

Service

Full Time +2
Part Time -2

Insurance

Carried +1
Not carried -1

Education

8 years +1
10 years +2
12 years +3
16 years +2

Clubs

Belongs to Clubs +1
Does not Belong -1

Experience

Previous Life Insurance Experience +1

Confidence

Replies to Question: "What Amount of Insurance Are You Confident
of Placing Each Month" +1
Does Not Reply -1

All of these significant items are the results of study of the personal history blanks. They do not represent guesses as to what ought to distinguish successful from unsuccessful agents, but items upon which successful and unsuccessful men actually have differed. Of course, in no single case will all these items be significant, various factors will operate in some instances to cause certain items to lose their predictive value for certain men. It is not upon any single item, however, that we will base our conclusions, but rather upon the total result from all the items.

The significance of the items may be explained as follows:

(1) *Education*. Through the high school period, schooling seems to be increasingly important to the life insurance salesman. When, however, we reach college, other factors seem to enter in to render the college man dissatisfied with this profession, and for this company, at least, the college man is not as successful as the high school graduate. It would be well here to mention the fact that the college group is a small one for this company; with a larger one our results might well be different. Even with the small group involved, there are several notable exceptions. If the college man goes into the field with the determination to succeed in life insurance salesmanship, and to make it his profession, there is no reason why, with his superior training, he should not go beyond the high school man. In the past, the difficulty has probably been chiefly one of attitude, the college man considered the field an experiment in which he might make some money while he was casting about for something to do, and

did not go in with the real desire to work, and to continue working. Hence the results in this study which is, of course, based on achievement in the past.

(2) *Age*. Up to age 40 at entry, there is a steadily increasing chance for the applicant to become a successful salesman, 30-40 being the most favorable period, inasmuch as by that time a man has amassed a large enough fund of general experience to allow him to see a situation through the eyes of a prospect and thus render the prospective buyer the best service. After age 40, he generally is too fixed in his ideas and habits to be able to make the necessary adjustments upon embarking upon a new line of work.

(3) *Occupation*. Whether an applicant has been previously engaged in an occupation in which he came in contact with people is significant as indicative of his probable success as a life insurance salesman. The fact of having worked with people, of having had to establish connections with them, renders the approach to a prospect relatively easy to the man who has been engaged in a social occupation, while to his fellow applicant who has been busied with individualistic work, it is an extremely difficult task. Examples of social occupations would be salesman, lawyer, social service worker, ex-army or navy man, etc., while of unsocial ones would be clerk, mechanic, chauffeur, etc.

(4) *Marital Status*. The married man with his increased responsibility rendering 'making good' more necessary is a better prospective agent than the single man. The man with dependents may be considered equivalent to the married man.

(5) *Insurance*. Whether the applicant carries insurance or not is significant as a factor in determining whether he will become a good agent, inasmuch as it indicates the fact that he is sold on the value of his own product. If he has not convinced himself of the value of life insurance, he is not likely to convince others.

(6) *Service*. The applicant who states that he intends to devote his full time to selling insurance is far more likely to do his utmost to make a success and to give the profession a fair trial. On the other hand, the man who makes life insurance salesmanship merely an occupation for his spare time whereby he can earn a little extra money, is not likely to become a good agent. He will be interested possibly in selling insurance, but not in seeing to it that the policy remains sold, and he is less likely, through his work, to be in himself an advertisement for his company.

(7) *Clubs*. The man who belongs to clubs and enjoys meeting others ought to be scored positively. He is more likely to be able to meet his prospects on their own ground and to sell them. This factor is closely connected with the sociability of the previous occupation.

(8) *Confidence*. The men who reply to the question "What amount of insurance are you confident of placing each month" tend to be better prospects for successful agents. Just what guess they hazard is not of importance. But the fact that they have confidence enough in their own ability to make some prophecy is important.

(9) *Life Insurance Experience*. Previous life insurance experience is of value to an applicant, although, of course, the reasons for his change from one company to another must be ascertained.

After the total score for each of the 502 personal history blanks had been made up, the first year's production record for each agent was secured. The blanks were then divided into three groups, (1) those who made a score below 4, (2) those who scored between 4 and 8, and (3) those who scored above 8. The lowest score actually made was —5 and the highest, 14. Each class was then divided into three parts, those whose production was below 50 units, those with a production between 50 and 100 units, and finally, those who were over 100 units. The number falling into each class was computed. The following chart indicates the results of this grouping. On the ordinate have been plotted the production units, and on the abscissa, the scores on the personal history blanks. The double line marks the critical score placed at 4.

There are 243 men who fall below 4 on the personal history blank, that is, 243 men who lie below the critical score indicated by the double line. Of the 243, 232 also fall below the minimum satisfactory production which has been placed at 50 units. Thus, of men scoring below 4, 96% are failures in production.

The cases to the right of the second vertical are agents who scored 8 or over on their blanks. Of these 68 men, 36 had a production of 50-100 units or over 100 units, and so, may be termed moderately to very successful. Therefore, of men scoring above 8 on the history blank, 53% are successful, as opposed to 4% of successes among those scoring below 4.

In between the first and second verticals are the men who made between 4 and 8 on their blanks. The results here are not as striking as in the other groups. In this division,

although 83% are failures, 17% are successes. Should other tests be given, they would probably serve to differentiate further among the agents forming this group.

Looking at the chart the other way, we find that of the 433 agents whose production was less than 50 units, or in other words, of those below the first or lower horizontal, 54% fall into the group whose score on the personal history blank is below 4. That is, this group of agents scoring below 4, the critical score, contains more failures in production than do the other two groups of those above 4 put together. If, on the other hand, we turn to the group which we may term

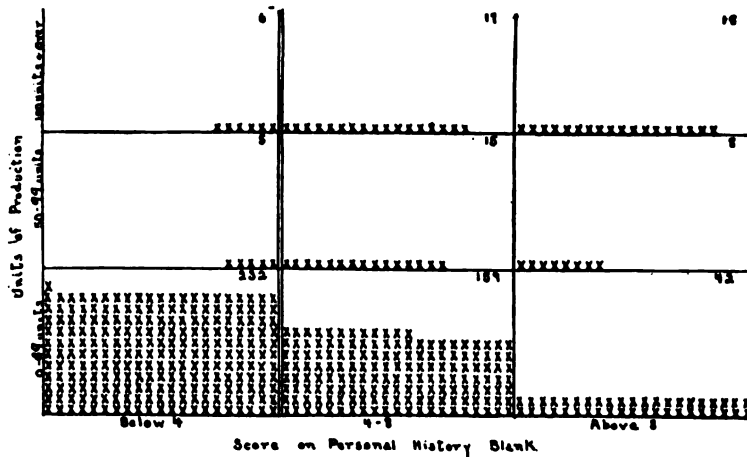


Chart Showing Relation of Personal History Score to
- Production Record

moderately to very successful, we find that of the 68 agents composing this group whose production is over 50 units, only 11, or 16%, will be found among those in the group scoring below 4 on the personal history blank.

It is evident then that the critical score should be set at 4. By cutting out those falling below that figure, we eliminate 54% of the subsequent failures, while 84% of the subsequent successes remain with the company. If the critical score were set higher up at 5 or 6, it is true that more failures would be eliminated, but at the same time more successes would be cut off,—an unprofitable operation.

The study of these 502 blanks has, therefore, indicated that for a life insurance company, the score on the personal history blank bears a positive relationship to the applicant's future success, and that on this blank a lower critical score may be set, below which it would not be worth while to license an applicant. The same would probably hold true for other sales forces in other occupations, but insofar as each industry presents a different problem, a different method of scoring would have to be worked out in each case.

THE SELECTION OF A SUCCESSFUL SECRETARY

A. T. POFFENBERGER, Columbia University

One of the high grade secretarial schools for women in New York City wished to try out a psychological examination as a means of selecting more competent and promising material and for raising the general standard of the school. The Army Alpha intelligence test was suggested and tried for one year. The data accumulated may be of service to the psychologist both by adding to what is already known about this examination, and also by indicating some of the difficulties that are met in applying it in such a practical situation.

The school is small enough so that each student is well known to the instructors and on that account the conditions are favorable for making estimates of fitness. Sixty-two students in the school were given the Alpha test. The school authorities supplied the following data:

1. A classification of the students into four groups as to their general school standing: Excellent, Very Good, Fair and Below the Standard of the School.

2. A classification of the students into ten groups according to their proficiency in each of their four most important studies, namely: Typewriting, Stenography, Accounting and English.

3. An order of merit arrangement according to proficiency in each of these studies.

4. A special report on certain students who were considered unsatisfactory.

Table 1 shows the coefficients of correlation (Spearman) between the order of the students in each of the four studies and in each of the eight parts of the Alpha test; also the coefficients of correlation between their order in each of the four subjects and that obtained from their total score in the Alpha test. If the ranks given for the four studies are averaged into what might be called a school rank, its correlation with the Alpha scores is represented by the coefficient $+.65$. Most of the correlations seem fairly large. The highest is between the test and English ($+.70$) and the lowest is between

the test and typing (+.46). The three coefficients in the body of the table that are lowest are those between typing and "Information" (+.11) and between accounting and "Oral Directions" (+.17) and "Practical Judgment" (+.10). The final coefficient of +.65 is about as high as has been obtained between college work and an intelligence test. One might then conclude that intelligence plays about as much part in secretarial school work as in general college work.

TABLE 1

	Oral Directions	Arithmetical Problems	Practical Judgment	Synonym- Antonym	Disarranged Sentences	No. Series Completion	Analogies	Information	Total Score
	1	2	3	4	5	6	7	8	
Typewriting.....	.59	.78	.45	.41	.46	.60	.62	.11	.46
Stenography.....	.60	.77	.39	.60	.64	.55	.66	.45	.55
Accounting.....	.17	.78	.10	.51	.60	.73	.40	.52	.50
English.....	.43	.63	.59	.71	.80	.71	.73	.63	.70

Judging from the size of the coefficients of correlation given in this table, certain of the separate parts of the Alpha examination are as effective for discovering typing and stenographic ability as are the special tests which have been tried out for that purpose. For example, the following coefficients were obtained by Rogers (unpublished) with two of his groups of skilled operators.

TABLE 2

	Typing	Stenography
Verb-object.....	.35	.36
Number checking.....	.39	.07
Color Naming.....	.38	.34
Agent-Action.....	.28	.23
Action-Agent.....	.26	.47
Substitution.....	.20	.40
Directions.....	.19	.46
Mixed Relations.....	.04	.31
Opposites.....	.06	.45

Without making any detailed comparison, it will be observed that most of the Alpha coefficients are higher than those in this group. It would seem to be advisable, when using the Alpha examination for testing special groups, to work with scores of the component tests as well as with the total score. Clues as to the measures of special capacity may thereby be obtained.

Table 3 gives some indication of the degree to which each of the eight parts of the Alpha examination is effective in picking the good from the poor in each of the four types of work, typing, stenography, accounting and English. The best ten and the poorest ten students in each of these subjects were selected, and their average scores in each of the eight parts of the Alpha examination were calculated. The average scores were reduced to percents of the maximum scores obtainable in the different parts, and these are given in the table. The differences between the records of the best and poorest ten are also given in the table.

TABLE 3

		Per cent of Maximum Score							
<i>Part of Alpha</i>		1	2	3	4	5	6	7	8
Typing	Best 10	78	59	81	86	88	63	74	73
	Poorest 10	54	34	66	72	80	44	47	68
	Difference	24	25	15	14	8	19	27	5
Stenography	Best 10	81	63	79	90	96	62	80	77
	Poorest 10	68	39	70	71	75	49	48	68
	Difference	13	24	9	19	21	13	32	9
Accounting	Best 10	81	62	71	88	93	66	84	75
	Poorest 10	70	39	70	74	76	45	62	63
	Difference	11	23	1	14	17	21	22	12
English	Best 10	83	63	77	92	96	68	86	80
	Poorest 10	75	51	64	68	71	43	50	59
	Difference	8	12	13	24	25	25	36	21

The school authorities in this case were interested in establishing a minimum score for entrance into their school. They were further interested in individuals rather than in statistical results. They wanted to know how their present students whose qualities were known would have been affected by establishing a minimum Alpha score at a certain point. To answer this question the students were distributed according to their appearance in the ten groupings provided by the instructors and according to their Alpha rating expressed in steps of ten points each. Such a distribution chart was prepared for each of the four studies and for a combination of these. The results were not satisfying. If the minimum score was set at such a level as to retain the good students (Alpha score 90) it did not eliminate the unsatisfactory stu-

dents; and if set high enough to eliminate the poor students, too many of the good students were eliminated. These tables will not be reproduced, as a condensed one will serve as an illustration.

Table 4 shows the distribution of students into four classes according to their general fitness as estimated by their instructors and into eight classes according to their score in Alpha. It will be seen that no one who made a score less than 90 made a satisfactory record in the school; also that no one who

TABLE 4
Alpha Score

School Standing	50-69	70-89	90-109	110-129	130-149	150-169	170-189	190-212
Excellent					4	13	7	1
Very Good			1	1	8	5	2	
Fair			1	7	3	3		
Below School Standing		1	2	1	2			

made a score above 171 made a record in the school worse than "Very Good." If the critical score were set at 90, only one of the six unsatisfactory students would be eliminated. If it were set high enough to include half of the unsatisfactory students, then two who were satisfactory would be eliminated. To eliminate all of the six unsatisfactory students would mean the elimination of twenty-five satisfactory students.

An examination of the cases where the students stood high in school work and low in the test shows clearly the importance of other than intelligence traits. Most important of these is a genuine interest and liking for the work and next to this stand loyalty, personal appearance, manner, steadiness and other social traits, which the test does not measure. One of the two cases of relatively high intelligence and low school standing, was a woman who, previously well-to-do, was forced by circumstances to earn her living. Lack of interest in secretarial work was obviously the difficulty here.

The writer believes that performance in an intelligence examination such as Alpha may not be an entirely adequate and economical basis for selection of secretaries any more than for mill workers and salesclerks. To choose out of all

available candidates those of the highest intelligence does not guarantee the selection of those having the character traits needed for successful secretarial work. Consideration should be given to both these aspects of the personality.

From the psychologist's point of view it might not be difficult to decide upon a critical intelligence score for the purposes of this school. For instance, a critical score of 130 would eliminate only 2 above average and 12 below average. But here as in most institutions where tests are being established the keenest interest centers upon the individual case. If one or a few good students be lost, the test may thereby be condemned. "If a man have a hundred sheep and one of them be gone astray, doth he not leave the ninety and nine, and goeth into the mountains, and seeketh that which is gone astray? And if so be that he find it, . . . he rejoiceth more of that sheep than of the ninety and nine which went not astray."

THE RELIABILITY OF JUDGMENT OF PERSONAL TRAITS

By JOHN SLAWSON, Division of Mental Defect and Delinquency, New York State Board of Charities

Two outstanding features in the science of mental measurement clearly indicate the necessity for the standardization of judgment evaluation of personal traits: (1), an unfortunate lack of objective measures for human traits other than intelligence, or educational attainment, so that evaluation of an individual in the shop, factory or school is still largely a matter of comparative judgment, such as ranking in order of merit; no contention is made that the judgment method is as desirable as the standardized scale method employed in evaluating intelligence, but since the former is used of necessity, it must be raised to as objective a level as possible; (2), in the derivation of a great many of our objective tests for intelligence and specific aptitude, we incorporate judgment evaluation in the composite criterion which determines the authenticity of these tests.

In this paper the writer intends to give a brief report of the results of a study¹ made in cooperation with the Bureau of Research of the Department of Education of the City of New York, the purpose of which was to determine the possibilities of the order of merit method in evaluating personal traits of individuals engaged in a specific profession or work. Eleven personal traits were chosen as defined below, and independent ratings were made by 31 judges distributed among six schools ($5+5+4+7+6+4=31$).

In each of the six schools, twenty-five teachers known to all of the judges chosen in that school were rated in the eleven traits, which were selected because of (a), supposed importance in the teaching profession, (b), competency of at least five raters to make judgments in them, and (c), distinctness or exclusiveness.² The judges were asked to rate the teachers

¹ The original report bearing the same title as this paper is on file at the Psychological Library of Columbia University, in manuscript form.

² Thorndike, E. L., "A Constant Error in Psychological Rating," *Jour. of App. Psychol.*, 1920, pp. 25-29.

in order of merit, i. e., placing the one who possessed most of the quality in question first and the one who possessed least of the quality last, by the use of the Stenquist "Card-list" Combination Ranking Card,³ which makes the task much less laborious to the rater and helps maintain uniformity in materials. Each judge, of course, made all ratings independently of the other judges.

- | | Trait | Definition |
|---------|--|--|
| No. 1. | <i>APPEARANCE.</i> | (Personal neatness in dress, cleanliness, etc.). |
| No. 2. | <i>TACT.</i> | (Ability to deal with others without giving offense). |
| No. 3. | <i>PUNCTUALITY.</i> | (Habit of being on time). |
| No. 4. | <i>EFFORT.</i> | (How hard does this person try?). |
| No. 5. | <i>JUDICIAL SENSE.</i> | (Fairmindedness, impartiality). |
| No. 6. | <i>LEADERSHIP.</i> | (Ability to lead, guide, direct, influence). |
| No. 7. | <i>COOPERATIVENESS.</i> | (Willingness to work effectively with others). |
| No. 8. | <i>PROFESSIONAL INTEREST AND GROWTH.</i> | (Interest in becoming a better teacher. Professional magazines, continued courses, lectures, associations, etc.). |
| No. 9. | <i>UNDERSTANDING OF CHILDREN.</i> | (Insight into child nature. Success in handling children). |
| No. 10. | <i>COUNTERACTING FACTORS.</i> | (This calls not for a judgment of the persons being ranked, but for a judgment of the <i>environment</i> in which each is working. Rank the person who is working under the greatest handicaps, No. 1, or first; the one working under the next greatest handicaps, 2d, etc.). Counteracting factors include such things as: especially difficult classes; poor equipment; depressing relations with other teachers or with supervising staff; bad hygienic conditions; poor health, etc.— <i>ALL AS COMPARED WITH OTHERS OF THE GROUP BEING JUDGED.</i> |
| No. 11. | <i>ALL-ROUND VALUE TO SERVICE.</i> | (As compared with all others. Not a <i>total</i> of the other items. A single estimate in toto disregarding specific items above). |

³ Stenquist, J. L., "An Improved Form of Rating by the Order of Merit Method." *Jour. Ed. Psychol.*, Dec. 1920, p. 526.

After an interval of about two weeks, the ratings were repeated by all of the judges independently of the first ratings. Sometime after the second set of ratings were made, the judges were asked to rate the same teachers using the same method as before, in acquaintance, i. e., placing the person best known to the rator first, and the one least known last. The judges were also requested to answer a questionnaire, for the purpose of ascertaining the qualitative aspects of the data, such as the kind of criteria used in making ratings, etc. Statistical treatment of the first set of ratings the second set, the ratings in acquaintance and the answers to the questionnaire enabled the writer to answer either provisionally or definitely the following questions:

1. Which of the traits chosen lend themselves best to objective evaluation?
2. Do the traits tend to retain their positions for relative objectivity in different groups?
3. To what degree can we reduce the differences in criteria employed?
4. What is the effect of acquaintance upon variability of trustworthiness of judgment?
5. What is the relation between judicial capacity and judicial consistency?
6. What is the effect of more than one trial of a rator's judgment upon judicial agreement?
7. What is the relation between judicial capacity and official position?
8. Does the distribution of judicial capacity tend to be specific or general? These questions will be dealt with separately in brief form. For a full account of statistical treatment employed and unabridged tables reference should be made to the original manuscript.⁴

1. Relative Objectivity of Personal Traits.

The degree of objectivity of a personal trait may be determined by ascertaining the degree of group agreement, that is, the greater agreement there is among competent judges in assigning positions to subjects (independently of each other), the more objective is the trait.⁵ Table I gives the average coefficients of group agreement of the eleven personal traits arranged in order of magnitude. Each average coefficient of agreement ($\text{Av. } \bar{r}_{pq}$) is the average of the average co-

⁴ *Ibid.*

⁵ Hollingworth, H. L., "Experimental Studies in Judgment." *Archives of Psychol.*, No. 29, pp. 116, 118.

efficients of correlation between any judge, p , and any other judge, q , in each of the six groups of judges. In other words, each of the \bar{r}_{pq} 's in Table I is the average of 68 inter-correlations. The P. E.'s are simply the variability coefficients for the six separate groups. The r_{pq} , the coefficient of agreement for each of the six groups of judges, not shown in Table I, was obtained by the use of the following formula devised by Professor T. L. Kelley, which makes possible the computation of a group correlational index directly without the need of computing the individual correlations between the judges:

$$\bar{r}_{pq} = 1 - \frac{M(2+4n)}{(M-1)(n-1)} + \frac{12\Sigma(SX)^2}{M(M-1)n(n^2-1)}$$

\bar{r}_{pq} is the same as Pearson's r .

TABLE I

Final Position	Name of Trait	Av. \bar{r}_{pq}	P. E. (t-o)
1	All-Round Value to Service.....	603	.032
2	Cooperativeness.....	522	.013
3	Leadership.....	503	.031
4	Effort.....	491	.026
5	Understanding of Children.....	472	.036
6	Professional Interest & Growth.....	470	.027
7	Appearance.....	460	.016
8	Tact.....	453	.033
9	Punctuality.....	408	.040
10	Judicial Sense.....	335	.027
11	Counteracting Factors.....	251	.038

Final Position of Traits with their average coefficients of agreement and unreliabilities of these averages.

Table I shows that "All-Round Value to Service" holds first place in objectivity; it has the highest coefficient of intra-group agreement. Judging from its P. E., its position in the series of eleven is relatively secure. It is to be noted that this is a complex or composite trait; the others are elemental or specific traits. This result, i. e., the greater objectivity of a composite trait as compared with specific traits is in accord with Wells' findings in his study of literary merit.⁶ "Counteracting Factors," the condition that tends to lower the efficiency of a teacher, retains the last position with about as high a

⁶ Wells, F. L., "A Statistical Study of Literary Merit," *Archives of Psychol.*, No. 7, p. 14.

reliability as "All-Round Value" retains first place. "Appearance" occupies a surprisingly low position. If we consider the fact that the correlation between two intelligence tests, one verbal and one non-verbal, is about .60, we can see the significance of the first three coefficients in Table I as regards objectivity.

2. Intergroup Agreement as to Relative Objectivity.

If we examine the differences between the coefficients together with the P. E.'s in Table I, we can see at a glance that the relative positions of the traits are insecure. This is, of course, due to the lack of a high degree of agreement between the six schools as regards the relative objectivity of these personal traits. Table II illustrates this fact more clearly. Here we see that the relative positions of the traits obtained from the coefficient of agreement in each school are by no means alike in each of the schools. Traits numbered 5, 7,

TABLE II

Trait No. Trait	Sch. I	Sch. II	Sch. III	Sch. IV	Sch. V	Sch. VI
1 Appearance.....	9	7	4	4.5	8	6.5
2 Tact.....	6	3	7	10	10	3.5
3 Punctuality.....	2	2	9	9	9	10
4 Effort.....	5	6	8	1	4	8
5 Judicial Sense.....	8	10	11	11	7	9
6 Leadership.....	3.5	9	3	6	2	5
7 Cooperativeness.....	3.5	5	1	4.5	5	3.5
8 Prof. Int. and Growth.....	10	8	2	7	3	6.5
9 Understanding of Children.....	7	4	10	2.5	6	1
10 Counteracting Factors.....	11	11	6	8	11	11
11 All-Round Value.....	1	1	5	2.5	1	2

Relative positions of traits obtained from the coefficients of agreements in each school (1 indicates highest \bar{r}_{pq} and 11 indicates lowest \bar{r}_{pq}).

10 and 11 hold their positions with a fair degree of consistency in that they at least stay in either the upper or lower half of the series of eleven. The positions of the rest of the traits are quite inconsistent. Trait No. 3, "Punctuality," occupying 2d position in the first two schools and 9th and 10th positions in the last four schools, shows indications that a variable operating in Schools I and II, in a manner opposed to that in Schools III, IV, V and VI is tending to make this trait objective in the former case, and subjective in the latter.

This disagreement with respect to objectivity among apparently homogeneous groups is of practical significance, for

should the intergroup disagreement be inherent in rating by the order of merit method, local standardizations for each group would become necessary, thus greatly diminishing the usefulness of rating. The trait that would lend itself best to objective evaluation in one school of a large school system, or in one department of a large industrial organization, would not do so in another school or in another department.

3. The Question of Criteria Employed.

Although personal interpretation by the judges cannot and possibly should not be entirely eliminated, qualitative analysis of the questionnaire which was sent to the judges showed that standardization of the environmental conditions under which judgments are made and of the elements to be included in the interpretations will greatly reduce the disagreement and inconsistencies referred to above. The following are some of the more important points noted in regard to the standardization of criteria.

A. Consulting Records: in schools I and II, practically all of the judges consulted records when rating for "Punctuality," but in the other four schools only two out of 21 judges consulted records. This explains the inconsistency referred to, which tended to increase intergroup disagreement, i. e., the trait occupied a high position for objectivity where records were consulted and a low one where they were not. This factor can of course be easily either eliminated or standardized. With the trait "Professional Interest and Growth," a somewhat similar situation occurred—records of professional courses and examinations taken were consulted by some judges and not by others.

B. There was an indication that conditions were more favorable in some schools than in others for making ratings in a given trait. It became evident upon reading the questionnaire answers that certain traits had greater judicial advantage in one school than in another, because one school placed greater emphasis upon the activities involving these traits than another school. In schools II and VI where the trait "Tact" held a high position with respect to judicial agreement, we find statements characteristic of these schools only, such as, "Most teachers in this school have some executive position which makes judgment in this trait fairly easy." In the trait "Effort," some schools utilized as criteria extra-school activities, and others intra-school activities. "Counter-acting Factors" suffered in practically all schools due to lack of personal acquaintance with the conditions under which the

teachers were working, and it may be for this reason that it occupies a consistently low position for group agreement. The data points quite conclusively to the fact that ratings made in a trait which is brought more vividly or prominently to the attention of raters during an entire year are subject to less intra-group disagreement than ratings made in a trait rarely brought to the attention of the raters (or playing a relatively unimportant rôle in the life of that particular school).

C. It appears that simple definitions of traits should be supplemented by specific items wherever possible, in order to reduce variability in criteria. In "Appearance," some allowed for financial condition of the persons rated; others emphasized how the wearing apparel suited the subject.

4. Effect of Acquaintance.

If we utilize the ratings made for acquaintance as heretofore described, and let

r_{1g} = the correlation between the rankings in merit given by judge I and the sum of the rankings given by the rest of the judges of his group *excluding* his own;

r_{1a} = the correlation between the rankings in merit given by judge I and his rankings in acquaintance;

r_{1g} = the correlation between the rankings in acquaintance given by judge I and the sum of the rankings in merit given by the rest of the judges of his group, *excluding* his own,

we can partial out the factor of acquaintance and get

$r_{1g.a}$ = the correlation between the rankings in merit given by judge I and the sum of the rankings given by the rest of the judges *excluding* his own, when acquaintance is kept constant (or when the judge in question is made equally acquainted with all the subjects rated).

By this process we determined to what extent *relative* acquaintance with the subjects rated contributed to the judge's agreement or disagreement with the rest of the judges. Does he deviate more from the group arrangement in rating the subjects whom he knows least, than in rating those subjects whom he knows most? It is impossible, due to lack of space, to give here the gains or losses when acquaintance is made constant for each of the 31 judges in each of the 11 traits. A summary is given in Table III by recording for each trait the average $r_{1g}'s$, i. e., the average judicial agreement when the factor of acquaintance is untouched and the average

$r_{I.g.a's}$, i. e., the average judicial agreement when acquaintance is made constant. In the fourth column are given the differences as indicated, and in the fifth the P. E.'s of these differences. A positive difference indicates, of course, that acquaintance was a handicap or hindered judicial agreement, while a negative difference shows that relative acquaintance promoted judicial agreement, because when acquaintance was made constant (ruled out) judicial agreement was decreased.

TABLE III

Trait No.	Av. $r_{I.g.a}$	Av. $r_{I.g.a}$	Diff. ($r_{I.g.a} - r_{I.g}$)	P. E. (diff.)
1	.602	.636	+.034	.026
2	.562	.583	+.021	.029
3	.558	.541	-.017	.034
4	.662	.648	-.014	.031
5	.494	.445	-.049	.034
6	.619	.583	-.036	.031
7	.642	.627	-.015	.024
8	.581	.574	-.007	.031
9	.592	.630	+.038	.036
10	.400	.385	-.015	.038
11	.661	.651	-.010	.024

Showing the influence of relative acquaintance in hindering or promoting judicial agreement in a given trait.

It will be noticed that although there are 8 negative differences, the unreliabilities are so large (exceeding the differences in 7 out of the 11 traits, and in the remaining 4 being slightly less than the differences), we can fairly safely conclude that the influence of the difference in relative degree of acquaintance on judicial agreement in situations such as were here encountered is negligible.

This inability of acquaintance to definitely either raise or lower judicial agreement becomes evident upon considering the several ways in which this factor may operate. For although lack of acquaintance with one or several subjects may result in chance ratings, thus lowering the correlation between the unacquainted judge, and the rest of the group of judges, intimate acquaintance between a rator and subjects may also lower the correlation between the intimately acquainted judge and the rest of the judges, by the exercise of prejudice either due to friendship or to the discovery of peculiarities in the subject which are particularly abhorrent to the rater (and probably unknown to the less intimately acquainted judges). The positive and negative influences would then in the long run tend to balance each other, that is, acquaintance would have little or no effect. This finding is, of course, no argument in favor of phrenology, because zero acquaintance was

eliminated before we started by choosing as judges those who had at least some acquaintance with the subjects. The discussion refers only to relative degree of acquaintance and not to a total lack. This result should also not be confused with what we said about the advantage that the familiar trait has over the non-familiar one as regards group agreement or objectivity—"trait acquaintance" is an important factor, but relative degree of personal acquaintance with subjects is not.

5. Judicial Capacity and Judicial Consistency.

Judicial capacity is determined by the degree of the judge's agreement with the competent group of judges of which he is a member.⁷ Judicial consistency is determined by the degree of agreement between two or more independent ratings made by a judge.⁷ No demonstrable relation was found between the former and the latter. This relation was investigated with the intention of ascertaining whether it was possible to determine capacity from consistency, since the latter can be determined so quickly and easily. The results were negative.

6. The Effect of More Than One Trial.

Contrary to expectations, we found that intercorrelating the average rank for two trials given by each judge with the average rank for two trials given by each of the other judges did not increase group agreement or objectivity. As a matter of fact, the latter was lowered in some of the traits. In rating by order of merit, it seems that increasing the sampling of measures (ratings) does not bring us nearer the truth, i. e., increase group agreement.

7. Judicial Capacity and Official Position.

Table IV enables us to approximate roughly the relation

TABLE IV

Official Position	No.	Position for Judicial Capacity						Total for 1st 3 pos.		Total for Pos. below the third	
		1		2		3					
		No.	%	No.	%	No.	%	No.	%	No.	%
Principals	4	1	25	2	50	1	25	4	100	0	00
Assis. Principals	10	3	30	3	30	2	20	8	80	2	20
Teachers	17	2	12	1	6	3	18	6	36	11	64

Showing the relation between official position and judicial capacity.

⁷ Hollingworth, H. L., op. cit., p. 109; "Vocational Psychology," pp. 158, 159, 167 (New York, Appleton, 1916).

Py B Thain

between official position and judicial capacity. In interpreting the table the percentages of principals, assistant principals and teachers occupying the first three positions in judicial capacity should be compared with the percentages occupying positions below the third. The superiority of the prinipals and assistant principals as judges is clearly seen. The assistant principals are better judges than the teachers, and the principals are better than the assistant principals. It is, of course, unsafe to make a positive conclusion from the small number of representatives that we have in each official position, but the indications are that executives and supervisors make better judges than associates or, stated differently, there is a positive correlation between official position and judicial capacity.

8. Distribution of Judicial Capacity.

The distribution of judicial capacity for each judge was investigated by plotting the number of times every judge occupied each designated position for judicial capacity in his group. In Fig. (a) below, we see that Judge A of School I, where there were 5 judges and, therefore, 5 possible positions for judicial capacity in each of the 11 traits, occupies the 5th or last position in 10 out of the 11 traits. He is, therefore, a generally poor judge; in Fig. (b) we see a generally good judge (but not so marked as the former).

Fig. (a) Position-Judge A-School I	Fig. (b) Position-Judge C-Sch. II	Fig. (c) Position-Judge D-Sch. IV
1st	1st *****	1st *
2nd	2nd *****	2nd ***
3d	3d *	3d *
4th *	4th	4th **
5th *****	5th	5th **
		6th *
		7th *

On the other hand, in Fig. (c) we see a judge occupying every possible position in his group (there were 7 judges in this school, and, therefore, 7 possible positions), from best to poorest. Figures (a) and (b) are illustrations of general judicial capacity, and (c) illustrates specific judicial capacity. If we plotted all the graphs here for each of the 31 judges, it would be seen that, roughly speaking, figures of type (a) and (b) are predominant, indicating that there is a greater tendency toward a general distribution of judicial capacity for each judge than toward a specific distribution. There is,

however, hardly strong enough evidence to conclude that if a person is a good judge in one trait he will be an equally good judge in another trait. There is an indication that he will. This relation is somewhat similar to the one found with intelligence—the preponderance of a general distribution of intelligence over a specific distribution.

COMPUTING INTERCORRELATIONS OF TESTS ON THE ADDING MACHINE

By HERBERT A. TOOPS, Institute of Educational Research, Teachers
College

Any scale lately constructed for mental, educational or vocational measurement must have both high reliability and high validity. If it is to lay claim to any particular merit, it must correlate well with a valid objective criterion, or composite variable consisting of independent, socially derived and socially acceptable, measures of the ability being tested. High validity may be secured by proper selection and scaling of the elements of the individual tests composing the scale, and by weighting the individual tests in such manner that the weighted composite shall correlate most highly with the criterion. In this latter process the intercorrelations of a number of tests are often computed. This is a strenuous task unless systematic methods are used, since, if there are n variables

including the criterion, there will be $\frac{n(n-1)}{2}$ correlations to be

computed. It is the purpose of this article to outline a procedure by which these can be rapidly calculated.

The procedure is based upon the formula,

$$r = \frac{N \cdot \sum XY - \sum X \cdot \sum Y}{\sqrt{N \cdot \sum X^2 - (\sum X)^2} \sqrt{N \cdot \sum Y^2 - (\sum Y)^2}} \quad (1)$$

in which X and Y are *gross measures of transmuted scores*, as described below. The speed possible with this formula depends upon: (a) the fact that these transmuted scores are numbers which are so small in magnitude that they may be readily multiplied mentally to obtain the XY products; (b) the use of a listing adding machine, the printed record from which supplies the required sums of Formula (1). The adding machine record affords a ready check upon the correctness of the work, and legibly prints all writing ordinarily required in making tally marks, raw scores, and the like, in other methods.

TABLE I
STANDARD GROUPING TABLE FOR DETERMINING TRANSMUTED SCORES FROM RAW SCORES WHICH ARE POSITIVE IN AMOUNT AND RANGE
UPWARDS FROM ZERO

Transmuted Score	Raw Scores Which Will Yield the Given Transmuted Scores of the Stub of the Table, When the Class Interval (I) is:												Transmuted Score
	1	2	3	4	5	6	7	8	9	10	11	12	
1	0	0-1	0-2	0-3	0-4	0-5	0-6	0-7	0-8	0-9	0-10	0-11	1
2	1	2-3	3-5	4-7	5-9	6-11	7-13	8-15	9-17	10-19	11-21	12-23	2
3	2	4-5	6-8	8-11	10-14	12-17	14-20	16-23	18-26	20-29	22-31	24-35	3
4	3	6-7	9-11	12-15	15-19	18-23	21-27	24-31	27-35	30-39	33-43	36-47	4
5	4	8-9	12-14	16-19	20-24	24-29	28-34	32-39	36-44	40-49	44-54	48-59	5
6	5	10-11	15-17	20-23	25-29	30-35	35-41	40-47	45-53	50-59	55-65	60-71	6
7	6	12-13	18-20	24-27	30-34	36-41	42-48	48-55	54-62	60-69	66-76	72-83	7
8	7	14-15	21-23	28-31	35-39	42-47	49-55	56-63	63-71	70-79	77-87	84-95	8
9	8	16-17	24-26	32-35	40-44	48-53	56-62	64-71	72-81	80-89	88-98	96-107	9
10	9	18-19	27-29	36-39	45-49	54-59	63-69	72-79	81-89	90-99	99-109	108-119	10
11	10	20-21	30-32	40-43	50-54	60-65	70-76	80-87	90-98	100-109	110-120	120-131	11
12	11	22-23	33-35	44-47	55-59	66-71	77-83	88-95	98-107	110-119	121-131	132-143	12
13	12	24-25	36-38	48-51	60-64	72-77	84-90	96-103	108-116	120-129	132-142	144-155	13
14	13	26-27	39-41	52-55	65-69	78-83	91-97	104-111	117-125	130-139	143-153	156-167	14
15	14	28-29	42-44	56-59	70-74	84-89	98-104	112-119	126-134	140-149	154-164	168-179	15
16	15	30-31	45-47	60-63	75-79	90-95	105-111	120-127	135-143	150-159	165-175	180-191	16
17	16	32-33	48-50	64-67	80-84	95-101	112-118	128-135	144-152	160-169	176-186	192-203	17
18	17	34-35	51-53	68-71	85-89	102-107	116-122	134-141	151-159	168-179	185-197	204-215	18
19	18	36-37	54-56	72-75	90-94	108-113	126-132	148-155	164-170	182-189	198-206	216-227	19
20	19	38-39	57-59	76-79	95-99	114-119	133-139	152-159	171-179	190-199	209-219	228-239	20

Directions:—Determine the range of the variable under consideration, Range=(highest score—lowest score + 1). Divide the range by 20, and raise the decimal of the quotient to the next higher integer for the class interval. Using the column of raw scores corresponding to that class interval in the heading, locate in the left or right margin the transmuted scores corresponding to the gross scores of the several subjects in the variable under consideration.

Example: (1). An age group ranges from 18 to 58, Range=(58—18+1)=41. Class interval=41/20=3. Use column I=3 for transmuted these ages into smaller classes.

Example: (2). The scores on a test range from 0 to 35. Range=(35—0+1)=36. Class interval=36/20=2. Use column I=2 for transmuted these scores into smaller classes.

The steps in order are as follows:

1. Tabulate the crude gross scores on the different tests in parallel columns, with a given subject's scores all on the same horizontal row, A, Fig. 1.

2. Determine the range of crude gross scores in each test. The range of a variable=(highest score made in the variable minus the lowest score made in the variable plus 1).

Now, for each test divide the range, thus determined, by 20, and raise the quotient to the next higher integer. This quotient will be the class interval, I , which indicates for a given variable which column of the Standard Grouping Table (Table 1) is to be used in transmuting the crude gross scores to class scores. These transmuted class scores are the classes under which a given crude score would be tabulated if grouped and tabulated into not over twenty classes. This rule for determining I , examples of which are given in footnotes to Table 1¹ insures that there will be *as many classes as possible*

¹ In practice, Table 1 is made out for the variables under consideration; that is, *Table 1 fits only those tests where zero scores occur*. Thus for Test 11, with no zero scores, but raw scores ranging from 41 to 82, say, the total range is 42; and $42/20$ gives $I=3$. In this case, one would not use the $I=3$ column of Table 1, but would construct a table:

"Grouping Table for Test 11 only, $I=3$.

Transmuted Score	Raw Score
1	39-41
2	42-44
3	45-47
4	48-50
.."

As a desirable convention, the columns of raw scores should always begin with a multiple of the class interval as above and as in Table 1.

The above table insures that the *maximum number of classes* will result. It is evident that the ($I=3$)-column of Table 1 cannot be used since it does not go high enough to contain the maximum score on Test 11.

Again, the crude scores on Test 19 range from 32 to 59. Here it would be *possible* to use $I=3$ column of Table 1, but this would give *only ten* classes. Here $I=2$, and the table to be constructed is:

"Grouping Table for Test 19 only, $I=2$.

Transmuted Score	Raw Score
1	32-33
2	34-35
3	36-37
4	38-39
.."

with the restriction that there shall be *not more than twenty* classes.

3. From Table 1, make out a cardboard stencil card (B. Fig. 1), of gross scores and transmuted classes for the variables as shown. A shows only the first three cases of the original data sheet; B shows only the first five of the total twenty lines of the stencil card; while C shows the transmuted scores of the three names of A, and one additional name. Note that the column spacings of B and C correspond

Crude Scores in Variable No.

A. Crude Data	Name	1	2	3	4	5	6	7	
	Aloe, John..	8	4	8	7	7	25	5	
	Beatty, Ray	14	17	52	13	15	79	28	
	Cole, Henry	10	6	14	3	8	23	20	

I=2 I=1 I=3 I=1 I=2 I=5 I=2

B. Cardboard Transmutation Table	Class								Class
	1	0-1	0	0-2	0	0-1	0-4	0-1	1
	2	2-3	1	3-5	1	2-3	5-9	2-3	2
	3	4-5	2	6-8	2	4-5	10-14	4-5	3
	4	6-7	3	9-11	3	6-7	15-19	6-7	4
	5	8-9	4	12-14	4	8-9	20-24	8-9	5

Transmuted Class Scores in Variable No.

C. Typewritten Transmuted Data	Name	1	2	3	4	5	6	7	
	Aloe, John..	5	5	3	8	4	6	3	
	Beatty, Ray.	8	18	18	14	8	16	15	
	Cole, Henry.	6	7	5	4	5	5	11	
	Dorsey, Tom	7	11	10	6	3	9	8	

Figure 1. Showing relative positions of Crude Data Sheet, Transmutation Table, and Transmuted Data Sheet in obtaining the transmuted scores. *A* is the crude data, large scores. *B* is the transmutation table, in which the columns are taken from the standard grouping table according to the range of the variable in the respective columns. *C* is the transmuted data sheet, on which the largest possible gross score is 20.

with A. In A, it is assumed that the range of variables is such that the class intervals are: $I_1=2$; $I_2=1$; $I_3=3$; $I_4=1$; $I_5=2$; $I_6=5$; $I_7=2$; and these columns are taken from the corresponding columns of Table 1 in constructing the stencil card transmutation table, B.

4. Lay B upon A, allowing "Aloe, John" scores only to be visible. The transmuted scores are read from B, variable by variable, and either copied directly beneath on the transmuted data sheet, C, or else read to an assistant who copies them onto the transmuted data sheet, C. Note that "Aloe, John" score of 8 on Variable 1, is found in the same column of B in the 8-9 group, and means a transmuted class score of 5; that his score of 4 on Variable 2, is found directly beneath on B to mean a transmuted class score of 5, etc.

5. Have an assistant independently make the transmutation, and thus check the correctness of C.

6. Typewrite C, making *two* carbon copies, in addition to the original typed copy, using relatively new carbon paper so that the carbon copies will be clear and distinct. Type the variable numbers as headings of the columns. *After each third name allow an extra blank horizontal row.*

7. Proofread C, making corrections of any errors on the carbon copies also.

8. Add up each of the several n columns on the listing adding machine imitating exactly the spacing of C, i. e., leaving a blank *after every third* entry. Label each column carefully with the variable number.

9. Check the work by placing the adding machine slip alongside the corresponding column of C. It will be found that the typewriter and adding machine, having the same size platens, or feed rolls, make the same spacings; hence *corresponding entries will be in opposition throughout* and may be readily checked. The adding machine sum 5 at the feet of the columns, if the work has been correctly performed, will be the ΣX 's or the ΣY 's of the formula, since the terms X and Y are merely relative. There will be n such columns.

10. Mentally compute the *squares of the entries* of C and enter them into the adding machine in groups of three followed by a space. There will be n such columns. Label each column carefully with the variable number squared, thus, $(1)^2$, $(2)^2$, etc.

11. Have an assistant check the squares after the manner

of procedure 9 above. The correct sums at the feet of the columns are the ΣX^2 's, or the ΣY^2 's.

12. Cut up both carbon copies of C into vertical columns, one variable to a column strip.

13. Place carbon strip, n , against variables $n-1$, $n-2$, $n-3$,1 of original typed C in turn, obtaining for each an adding machine record of the XY products in groups of three entries and a space. Then place the $(n-1)$ carbon strip against the $n-2$, $n-3$,1 columns of C in turn in similar manner; and so on until *every column has been paired with every other column*. The multiplications can be done mentally in most cases. The largest possible product is $20 \times 20 = 400$. A simple multiplication table, constructed by the author, enables one at a glance to note any desired product not remembered, without the inaccuracies attendant upon coordinate locating, which is the fault of most multiplication tables. There are $\frac{n(n-1)}{2}$ such column multiplications. Label each slip carefully with the product of the two variables concerned; thus, $(1) \times (2)$, $(1) \times (3)$, $(1) \times (4)$, etc.

14. Have an assistant check the multiplications using the second carbon sheet strips. This work may proceed at the same time that the adding machine records are being made.

15. All the needed quantities are now known for substituting $\frac{n(n-1)}{2}$ times in formula (1) to obtain the several $\frac{n(n-1)}{2}$ intercorrelations. This may probably best be done in systematic form as shown in Figure 2. In this table,

15(a). The first entry of six in each large compartment is the ΣXY product derived from the checked adding machine records of procedure 14, for the variables at the intersection of whose coordinates the compartment is located. Notice that *these entries check for symmetry* in rows and columns about the diagonals.

15(b). The second entry of a large compartment is $N \cdot \Sigma XY$. A calculating machine may be used to multiply all ΣXY product-sums of the entire table in turn by the constant multiplier, N .

15(c). The third entry of a compartment is $\Sigma X \cdot \Sigma Y$, for the compartment considered; i. e., the ΣX and ΣY which are coordinates of the column and row respectively in which a given compartment is located. (The ΣX , ΣY , and ΣX^2 quanti-

ties may be written into their appropriate places in procedure 15a.) These products check for symmetry of rows and columns about the diagonals.

FIG. 2

The Systematic Solution of Equation 1 in a Table of Intercorrelations.
N= 81 persons. Variable 1 is a criterion of general business ability; variables 9, 11, 13 are tests.

$\Sigma X \rightarrow$		904	818	1096	622	$1/\sqrt{K}=1/\sqrt{H}$
ΣY	Variable No.	I*	9	11	13	
904	I	1.0000	$\Sigma XY = 9390$ $N \cdot \Sigma XY = 760590$ $\Sigma X \cdot \Sigma Y = 739472$ $L = 21118$ $L \cdot \frac{1}{\sqrt{K}} = 94.0148$ $r = .3811$	12391 1003671 990784 12887	7156 579636 562288 17348	.0044518815
818	9	9390 760590 739472 21118 85.6110 .3811	1.0000	11185 905985 896528 9457 38.3381 .2168	6363 515403 508796 6607 26.7844 .1182	.0040539357
1096	11	12391 1003671 990784 12887 72.8649 .3244	11185 905985 896528 9457 53.4712 .2168	1.0000	8553 692793 681712 11081 62.6535 .2766	.0056541408
622	13	7156 579636 562288 17348 76.5739 .3409	6363 515403 508796 6607 29.1632 .1182	8553 692793 681712 11081 48.9114 .2766	1.0000	.0044139895
A ΣX^2		=10712	9012	15216	5410	
B $N \cdot \Sigma X^2$		=867672	729972	1232496	438210	
C $(\Sigma X)^2$		=817216	669124	1201216	386884	
D $H = K = N \cdot \Sigma X^2$						
E $-(\Sigma X)^2$		=50456	60848	31280	51326	
F $\sqrt{H} = \sqrt{K}$		224.62413	246.67387	176.86153	226.55242	
F $1/\sqrt{H} =$.0044518815	.0040539357	.0056541408	.0044139895	

Row 1 of each compartment is ΣXY
Row 2 of each compartment is $N \cdot \Sigma XY$
Row 3 of each compartment is $\Sigma X \cdot \Sigma Y$
Row 4 of each compartment is $N \cdot \Sigma XY - \Sigma X \cdot \Sigma Y = L$
Row 5 of each compartment is $L \cdot 1/\sqrt{K}$
Row 6 of each compartment is $L \cdot 1/\sqrt{K} \cdot 1/\sqrt{H} = r$

*NOTE: Do not confuse this I (criterion variable) with I (class interval) of Table 1 and of formula (4).

15(d). The fourth entry of a large compartment is the second entry minus the third, or $N \cdot \Sigma XY - \Sigma X \cdot \Sigma Y$; that is, the numerator of the correlation coefficient whose solution is being sought in that compartment. These remainders also check for symmetry of rows and columns about the diagonals.

15(e). The next step is the simplification of the denominators, which is done in the last six horizontal rows of the table.

(A). The first of these six entries is ΣX^2 obtained from procedure 11.

(B). The second row is $N \cdot \Sigma X^2$.

(C). The third is $(\Sigma X)^2$, obtained by squaring ΣX of the same column, found at the top of the page.

(D). The fourth is $(B) - (C)$; or, $N \cdot \Sigma X^2 - (\Sigma X)^2$.

(E). The fifth is the square root of (D).

(F). The sixth is the reciprocal of (E), or

$$1/\sqrt{N \cdot \Sigma X^2 - (\Sigma X)^2}.$$

These reciprocals allow one to solve Formula 1, by a multiplication process (where a calculating machine is available). Thus, r is equal to the simplified numerator (an integral number derived in 15d) divided by the square root of the product of the simplified denominators of the two variables concerned (derived in 15(e) (D)); r is likewise equal to the simplified numerator, L , times the 15(e) (F) reciprocal of X times the 15(e) (F) reciprocal of Y . If we let:

L =simplified integral numerator of 15(d) procedure;

H =simplified integral denominator expression of the X -variable in the 15(e) (D) procedure;

K =simplified integral denominator expression of the Y -variable in the 15(e) (D) procedure.

$1/\sqrt{H}$ and $1/\sqrt{K}$ =the results of 15(e) (F) procedure for the X and Y variables respectively, entering into a given correlation r_{xy} .

$$\text{Then, } r_{xy} = \frac{L}{\sqrt{H} \sqrt{K}} = \frac{L}{\sqrt{H \cdot K}} \quad (2)$$

in which L , and H and K and $H \cdot K$ are *all integral quantities*, the solution of the equation being readily performed on a slide rule.

Or, by the second procedure, with a calculating machine,

$$r_{xy} = \frac{L}{\sqrt{H}} \frac{x}{1} \frac{1}{\sqrt{K}} \frac{x}{1}$$

in which L is an integral number and

$\frac{1}{\sqrt{H}}$ and $\frac{1}{\sqrt{K}}$ are decimal multipliers, as shown in Fig. 2.

15(f). Assuming a calculating machine available so that the solution may be effected by Formula (3), the results of procedure 15(e) (F) are now to be written on the right hand side of the chart, noting that like reciprocals belong to the same variable numbers. (In all this work, it is to be remem-

bered that the terms X and Y are purely relative; that is, that in the correlation r_{12} , variable 1 is X and variable 2 is Y ; in the correlation r_{23} , variable 2 is X and variable 3 is Y ; or that, finally, X is synonymous with "variable of the two mentioned first" while Y is synonymous with "variable of the two mentioned last." It follows that $\Sigma X_1 = \Sigma Y_1$, $\Sigma X_1^2 = \Sigma Y_1^2$, $H_1 = K_1$, etc.)

15(g). Now multiply the fourth entry of each large compartment in each given row by the 15(e) (F) entry at the end of the respective rows. On a calculating machine this 15(e) (F) entry, $1/\sqrt{K}$, is a constant multiplier for all compartments of a given row. These products do not check for symmetry.

15(h). Multiply *in columns*, each fifth entry of a large compartment by the 15(e) (F) entry, $1/\sqrt{H}$, at the foot of the respective columns. The result is r . This $1/\sqrt{H}$ entry is a constant multiplier for all compartments of a given column. The products must check for symmetry of rows and columns about the diagonals.

With a calculating machine one may omit recording procedures 15(b) and 15(c), and 15(e) (B) and 15(e) (C) and arrive directly at the integral quantities of procedures 15(d) and 15(e) (D). If accuracy of r to not more than three decimal places is required, one may omit 15(e) (E) and 15(e) (F); in that case, the correlation coefficient is given by formula (2), which is solved by slide rule.

It will be noted that 15(d) is the simplified numerator, *in integral numbers*; and that 15(e) (D) is the simplified denominator *in integral numbers* for a given correlation.

In the work of procedures 13 and 14, the carbon slips may be held in proper place by thumb-tacking sheet C and carbon slip to a soft pine work board. It will be found that the grouping of records in groups of three's will very greatly facilitate checking, eliminate checking errors, and relieve eye strain.

The numerous checks for symmetry in the solution of equation (1) in procedure 15 using Fig. 2 table, will be found to be a very adequate check on the solution of the equation. The final r 's, if the work is correctly done, will check for signs, decimals and absolute magnitude. The subtractions in procedure 15 (d) are algebraic; consequently r automatically comes out with the right sign if the work is done correctly.

With this method, the author's best record thus far is the production of the XY product sums of 20 correlations of 151

cases each in an average time of 6 minutes 23 seconds each. The checking showed 3 entries of the 3020 XY entries to be in error. These were readily detected and corrected. It should be noted by way of caution that even a slight error in ΣXY will make a large error in r ; with reasonable care, and a possible rechecking of the adding machine slips, no errors should escape. The computation speed outweighs the consideration that a possible error will absolutely invalidate a given r ; *statistical work should be accurate*. The computation of the XY products for $N=50$ persons will take a little more than 2 minutes per correlation; for $N=75$, a little more than 3 minutes, etc. Not so much time is required for checking; and finally, the systematic tabular form of solution of equation (1) is done more rapidly and with fewer errors than in any other method of solution; any errors in the solution of the equation will be located by the symmetry check as it is highly improbable that two errors will be made and of such an amount that the final r 's would check. If $r_{14} \neq r_{41}$, and if $r_{18} \neq r_{81}$, the errors may be quickly detected.

The greater the number of variables, the greater the saving in time effected by this method, as will be seen from the following tabulation of the total number of adding machine records needed, and the number of adding machine records per intercorrelation for $n=2$, $n=3$, etc., variables, as n grows larger.

TABLE 2

Table Showing the Total Number of Adding Machine Records Required, the Number of Intercorrelations Resulting, and the Number of Records per Intercorrelation, as the Number of Variables, n , Increases

n	Total No. of Records $= n^2 + 3n$	No. of Inter- correlations $= n(n-1)$	No. of Records per Intercorrelation $= \frac{n+3}{n-1}$
	$\frac{\quad}{2}$	$\frac{\quad}{2}$	$\frac{\quad}{n-1}$
2	5	1	5.0
3	9	3	3.0
4	14	6	2.3
5	20	10	2.0
6	27	15	1.8
7	35	21	1.7
8	44	28	1.6
9	54	36	1.5
10	65	45	1.4

The fourteen adding machine printed records required for 4 variables are shown in Table 3. Remember that the first four columns are adding machine duplications of the same

columns in Table C, which was previously derived from larger gross crude scores. None of the records of Table 3 are ever written down longhand, being instead always computed mentally or by simple multiplication table up to products of 20x20, and entered directly into the adding machine. This gains speed over the slow writing of the human machine and secures legibility as well. All multiplications up to 15x15 may be easily memorized; beyond this point a simple multiplication table will refresh the memory for products up to 20x20. Every test worker knows most of the squares up to (20)².

It remains yet to be pointed out how the true averages and standard deviations may be derived.

The standard deviation presents no problem.

The formulae are:

$$\sigma_x = \frac{I_X \sqrt{H}}{N} = \frac{I_X \sqrt{N \sum X^2 - (\sum X)^2}}{N} \quad (4)$$

$$\sigma_y = \frac{I_Y \sqrt{K}}{N} = \frac{I_Y \sqrt{N \sum Y^2 - (\sum Y)^2}}{N} \quad (4a)$$

In which I is the class interval, determined in Procedure 2, and H and K have already been determined in Procedure 15(e) (D).

The averages present a little more difficulty. The formulae for the averages, in terms of transmuted class scores, are:

$$M_x = \frac{\sum X}{N} \quad (5)$$

$$M_y = \frac{\sum Y}{N} \quad (5a)$$

These must now be located with reference to the *raw* scores. Suppose that by formula (5), $M_x = 3.47$ when the class interval of variable X is, $I_x = 3$. Our problem is thus to interpolate, in Table 1 in the ($I=3$)—column, a *transmuted class score* of 3.47 into a *raw* score. Now the *third* transmuted class is 6-8 in terms of raw scores, the *face value* of the class being 7.0.* We thus know that the *raw score* average is somewhat more than 7.0; it is really 47/100 of a class interval more than 7.0. Now, $(.47)(3) = 1.41$. And $7.0 + 1.41 = 8.41 = M_x$, in terms of *raw scores*.

Again, if by formula (5a), $M_y = 9.80$, when $I_y = 4$; then, M_y , in the terms of *raw scores*, $= 33.5 + (.80)4 = 36.7$.

The values of r obtained by this method are the identical values of r which would be derived from Fig. 1, C table, or from any exact method wherein the raw scores are grouped into coarser classes. By the procedure used in determining I, the grouping error of r , inherent in all grouping methods, will seldom, if ever, be more than one per cent of the magnitude of the correlation coefficient, since the coarsest possible grouping, by the schema here presented, is in the case of a range of 21, in which the gross score of 20 becomes the

TABLE 3

14 Adding Machine Records Required in the Solution of the 6 Intercorrelations of 4 Variables.

Columns 1, 2, 3, 4 are copied directly from the transmuted class scores, Fig. 1, C. Each column of this table represents a separate adding machine record.)

X_1	X_2	X_3	X_4	X_1^2	X_2^2	X_3^2	X_4^2	X_1X_2	X_1X_3	X_1X_4	X_2X_3	X_2X_4	X_3X_4
1	3	2	1	1	9	4	1	3	2	1	6	3	2
3	2	5	2	9	4	25	4	6	15	6	10	4	10
8	12	9	5	64	144	81	25	96	72	40	108	60	45
4	3	6	8	16	9	36	64	12	24	32	18	24	48
6	7	8	8	36	49	64	64	42	48	48	56	56	64
9	8	10	9	81	64	100	81	72	90	81	80	72	90
12	10	8	6	144	100	64	36	120	96	72	80	60	48
10	11	12	10	100	121	144	100	110	120	100	132	110	120
13	12	10	9	169	144	100	81	156	130	117	120	108	90
8	6	10	11	64	36	100	121	48	80	88	60	66	110
74	74	80	69	684	680	718	577	665	677	585	670	563	627
ΣX_1	ΣX_2	ΣX_3	ΣX_4	ΣX_1^2	ΣX_2^2	ΣX_3^2	ΣX_4^2	ΣX_1X_2	ΣX_1X_3	ΣX_1X_4	ΣX_2X_3	ΣX_2X_4	ΣX_3X_4

eleventh class. Since the correction, due to coarseness of grouping, becomes negligible if at least twelve classes are used, it will be seen that the denominator, 20, used in determining the class intervals, is a happy compromise between ease of mental multiplication and requirements for disregarding the grouping error involved by this method. The grouping errors in the use of this method may be entirely disregarded.

NOTE.—There are two current usages with respect to the face value of a class, 6-8. The first is that advocated by Pearson, involving the

determination that a score, 6, is *closer to 6* than to any other value. Consequently the class, 6-8, will include all gross scores of 5.500 up to and inclusive of 8.499; whereupon the face value of the class is $\frac{5.500 + 8.499}{2} = 7.0$. The second usage, growing out of the test con-

sideration that one cannot obtain a test score of 6 unless his ability is *at least 6*, considers the class, 6-8, to extend from 6.000 to 8.999, whereupon the face value of the class is $\frac{6.000 + 8.999}{2} = 7.5$. It is essential

to note that people advocating the latter procedure, when finding the average of a number of test scores *without grouping* are likely to add all together on the adding machine and divide by N without thinking to add the correction of $\frac{1}{2}$; when this error is made, the implicit assumption is of course that a score of 6 is *nearer 6* than any other figure, and this procedure is the first method referred to above. Consequently many errors occur in the second method. Inasmuch as the test units are not absolute measures from a stationary zero, such as the units of a yardstick, no possible good can come from adding the $\frac{1}{2}$ gross score correction, and confusion is the only likely outcome of the introduction of a procedure which is not universally applicable to statistics in other fields of research.

As a desirable convention, the lowest class of a contingency table might always be given in the form of the extreme limits classifiable in that class, after the fashion of some of the English writers. The face value of any given class can always be easily determined by taking one-half of the sum of its two extreme limits.

A GLIMPSE OF HIGH SCHOOL COURSES AS MEASURED BY THE OTIS TEST

By RUTH SWAN CLARK, Ph. D., Bureau of Vocational Service for
Juniors, New York City

Feeling the demands of modern industrial and social conditions, High Schools of today are offering special courses of study in order to train their students for various kinds of vocations. It is important to study, not only the subjects taught in these different courses to meet these industrial demands, but also to study the students who voluntarily choose to take these courses, for the purpose of discovering whether or not they show distinguishing mental characteristics.

It is with this problem in mind that the present survey has been made. The Otis Group Intelligence Test, Advanced, was given to students in both the entering and graduating classes of various High School Courses in New York City. In this way, studies were made not only of the students who voluntarily selected these courses, but also of those who were successful in meeting the demands of the courses, so that they were allowed to graduate. The High School courses which were chosen for this survey were the Academic, the Commercial, the Technical, the Industrial Arts, and the Dress-making.

The outstanding mental traits of the High School student of the Academic Course, so far as these could be detected by the use of the group intelligence test, are given in the following table:

The above facts indicate that the students who voluntarily chose the High School Academic Course were those of a general intelligence above the average for boys and girls of their own age. The range of intelligence within the group, however, was broad, so that one out of every 20 of these students were below the average, while nearly two-thirds of the whole class were above the average. These students were, moreover, accurate workers, having more than two-thirds of the work that they attempted correct. They also showed the ability to readily solve problems dealing with concrete situations. This ability was outstanding as compared with their memory for words in logical sequence.

TABLE 1

High School Academic Course. (Boys and Girls)

	1st Term	8th Term
¹ No. of Students.....	304	151
Av. Otis Score.....	129.9	166.3
² Av. Physical Age.....	14-2	16-0
³ Av. Mental Age.....	16-4	18-4
⁴ Av. Intelligence Quotient.....	115.2	114.5
Middle 50% Range of I. Q.....	98-124	110-118
Speed (% of Questions Answered).....	75.7%	77.4%
Accuracy (% of Questions Correct).....	70.2%	79.9%
⁵ Distinguishing Best Test.....	Arithmetic	Analogies
Distinguishing Poorest Test.....	Memory	Memory
No. of Students with Intelligence above average (I.Q. +110).....	63.1%	79.4%
No. of Students with Intelligence below Average (I.Q. -90).....	5.5%	.66%

The graduates of this course were of similar intelligence. The group was, however, more homogeneous; those with intelligence below the average were not with those who had completed the demands of the course. In fact, four-fifths of the class were of intelligence above the average for those of their age. The graduates were also more accurate workers, as compared with the entering students. Their memories again showed poorly in comparison with their other mental traits. In the case of the graduates, however, their ability to perceive relationships was superior to their ability to solve problems of a concrete nature. In general it may be said that this academic course adapted itself to students of a general intelligence above the average. A student with intelligence below the average would have but one chance in 150 to graduate, while the student of average intelligence would have one chance in 5. The correlation between the intelligence quotients of these graduates and the average of their school marks

¹ This number represents the total number of students enrolled in the course.

² The highest physical age credited to a student was 16 years.

³ This average was secured from the average score of the group translated into terms of the Binet Mental Age, as given in Otis Manual of Directions, 1921.

⁴ This quotient was secured by dividing the average mental age by the average physical age.

⁵ The "Best" and "Poorest" Tests were determined by a study of the number of errors as compared with the number of right answers in each test of each student. These tests were then arranged in order of their ease and difficulty for all students in all courses. Any deviation in this order shown by the students in any one course was considered a "distinguishing best" or "poorest" test.

HIGH SCHOOL COURSES MEASURED BY OTIS TEST 187

for the four years was .43 (.04 P. E.) (Pearson). This is not a high correlation, but in relation to similar correlations with the students in the other courses, it is the highest of all.¹

A similar study of the students of the Commercial Course shows the following facts:

TABLE II

High School Commercial Course. (Boys and Girls)		
	1st Term	6th Term
No. of Students.....	287	102
Av. Otis Score.....	119.1	147.9
Av. Physical Age.....	14-2	15-11
Av. Mental Age.....	15-7	17-5
Av. Intelligence Quotient.....	110.	109.4
Middle 50% Range of I.Q.....	98-117	103-114
Speed (% of Questions Answered).....	75.8%	80.7%
Accuracy (% of Questions Correct).....	64.4%	75.1%
Distinguishing Best Test.....	Memory	Proverbs
Distinguishing Poorest Test.....	Fol. Directions	Similarities
No. of Students with Intelligence above Average (I.Q. +110).....	43.2%	45%
No. of Students with Intelligence below Average (I.Q. -90).....	11.4%	2.9%

From the above facts, it appears that those who voluntarily chose the Commercial Course were students of high average intelligence. They were rapid in their work without being very accurate. They showed a better memory for words in logical sequence than the ability to adapt themselves to the following of directions of a specific nature. Those who graduated from this course, on the other hand, were quicker, more accurate, and more intelligent. It is true that the average intelligence of the graduating class was the same as that of the entering class, but there were fewer of those with average and less than average intelligence amongst the graduates than amongst those who were entering the course. It is interesting to note also that the graduates of this course showed more conspicuously their ability to grasp and to deal with meanings of words, as in the Proverbs Test, than they did their ability in memory: the trait which characterized the entering group. The correlation between the intelligence quotients of the graduates in this course and their average marks for their three years of work was .34 (.06 P. E.). This seems to indi-

¹ Cf. Proctor, W. M. *Psych. Tests and Guidance of H. S. Pupils. Jour. of Ed. Research*, Mon. I, 1921, p. 16.

² This course was for three years only.

cate that not as much dependence is placed in this course upon general intelligence, as in the Academic Course.¹

The following table shows similar mental traits for the students in a Technical High School Course:

TABLE III

Technical High School Course. (Boys)

	1st Term	8th Term
No. of Students	100	100
Av. Otis Score	121.9	147.3
Av. Physical Age	14-3	16-0
Av. Mental Age	15-8	17-5
Av. Intelligence Quotient	110.	108.9
Middle 50% Range of I.Q.	102-121	102-114
Speed (% of Questions Answered)	71.1%	77.3%
Accuracy (% If Questions Correct)	70.2%	78. %
Distinguishing Best Test	Opposites	Arithmetic
Distinguishing Poorest Test	Geometric Figures	Similari- ties.
No. of Students with Intelligence above Average (I.Q. +110)	50.9%	44.4%
No. of Students with Intelligence below Average (I.Q. -90)	10.5%	2. %

The boys who were just beginning this technical course seem to have been those of high average intelligence, not as quick in answering the questions as those of the Academic and Commercial Courses, but resembling the students of the Academic Course in their high degree of accuracy. They were conspicuous in their ability to note logical relations, as in the Opposites Test, but they less readily adapted themselves to following the detailed directions of the Geometric Figure Test. The graduates of this course were of similar intelligence, but were quicker and more accurate in their work. They showed themselves better able to deal with arithmetic problems, which were concerned with concrete situations, than with the forming of abstractions, which were called for in the Similarities Test. The correlation between the intelligence quotients of these graduates and their average school marks for the four years was .34 (.07 P. E.). This correlation resembles that found for the students of the Commercial Course more than that found for those of the Academic Course.

¹ Cf. Bock, W. F., *Intelligence of High School Seniors*, p. 145. Madsen, I. N., *Group Intelligence Tests as a means of prognosis in High School*. J. of Ed. Research, III., pp. 43-52; 396-398.

² Not all the entering and graduating students enrolled in these classes were tested. Instead, 100 students taken at random from each group were used to represent them.

HIGH SCHOOL COURSES MEASURED BY OTIS TEST 189

In Table IV are found the measurement of the same traits for the students of an Industrial Arts Course.

TABLE IV

High School Industrial Arts Course. (Girls)

	1st Term	6th Term
No. of Students.....	110	34
Av. Otis Score.....	125.8	132.7
Av. Physical Age.....	14-8	16-0
Av. Mental Age.....	16-0	16-6
Av. Intelligence Quotient.....	109.1	103.1
Middle 50% Range of I.Q.....	99-117	97-108
Speed (% of Questions Answered).....	74.7%	75.6%
Accuracy (% of Questions Correct).....	68.9%	71.9%
Distinguishing Best Test.....	Memory	Proverbs
Distinguishing Poorest Test.....	Analogies	Arithmetic
No. of Students with Intelligence.....		
above Average (I.Q. +110).....	44%	20.5%
No. of Students with Intelligence below		
Average (I.Q. -90).....	6.4%	8.8%

The students entering this Industrial Arts Course resemble more the students of the Commercial Course than those of the other courses. They are of high average intelligence, quick at their work without being especially accurate, and showing an ability in their memory for words in logical sequence. Their "poorest" test, however, was different. The students of the Industrial Arts course were less apt in perceiving relationships as such (Analogies Test) than they were to follow specific directions, which was the greater difficulty for the Commercial Course students.

In comparison with the students just beginning the Industrial Arts Course, the graduates of this course showed a similar speed in their work, but they were a little more accurate. As a group, they showed less general intelligence. This may be due to the fact that more intelligent students find the course progressing too slowly for them, or it may mean that the course adapts itself more to students with a special ability in art than to students with merely a general intelligence. The correlation between the intelligence quotients of these graduates and the average of their school marks, .36 (.10 P. E.), shows as close a relationship between intelligence and success in the course as has been shown in both the Commercial and Technical Courses. However, the size of the Probable Error of the correlation shows its unreliability.

The mental traits of the students of the Dressmaking Course are given in the following table:

TABLE V.
High School Dressmaking Course. (Girls)

	1st Term	6th Term
No. of Students.....	110	51
Av. Otis Score.....	109.6	119.8
Av. Physical Age.....	14-7	16-0
Av. Mental Age.....	14-9	15-7
Av. Intelligence Quotient.....	101.1	97.4
Middle 50% Range of I.Q.....	91-119	91-102
Speed (% of Questions Answered).....	70.5%	69.4%
Accuracy (% of Questions Correct).....	63.6%	70.7%
Distinguishing Best Test.....	Following Directions	Arithmetic
Distinguishing Poorest Test.....	Geometric Figures	Analogies
No. of Students of Intelligence above Average (I.Q. +110).....	26.3%	11.7%
No. of Students of Intelligence below Average (I.Q. -90).....	22.7%	21.5%

The girls voluntarily choosing this course in dressmaking were those of average intelligence. They were not as quick as the Commercial Course students but they showed the same degree of accuracy. Unlike the Commercial Course student, however, they displayed more ability in following directions accurately. They resembled the Technical Course students in the difficulty they found with the Geometrical Figure Test.

The graduates of this course were also of average intelligence. This group was, however, a more homogeneous one, for the entering students included many students of intelligence above the average for those of their age, while these did not appear amongst the graduates to such an extent. Again one might conclude that the course did not progress fast enough for students of higher general intelligence, or that the course adapted itself more to students with a special ability in hand work than to students with merely a general intelligence. The correlation between the intelligence quotients of these graduates and the average of their school marks for the three years was .35 (.06 P. E.). This correlation indicates as close a relationship between intelligence and success in the course as has been pointed out in the Commercial, the Technical, and the Industrial Arts Courses. Selection within these courses, therefore, does not take place on the basis of intelligence, as much as is apparent in the Academic Course, but

seems to depend upon other factors, such as ability in hand work.

In conclusion one may state that the High School Academic Course adapts itself to students with intelligence above the average for those of their age. Moreover, selection within the course seems to be based upon the possession of general intelligence. On the other hand, while special courses, such as the Commercial, the Technical, the Industrial Arts, and the Dressmaking require general intelligence of an average and high average degree, selection within these courses does not seem to be based upon this factor alone. In general, it is apparent that High Schools are not adapting their courses to train students of a general intelligence below the average. Such a student has one chance in 150 to graduate from an Academic Course, but three times as many chances to graduate from a Technical Course and five times as many chances to graduate from a Commercial Course.

No attempt has been made in this survey to study all the different High School Courses, nor yet to claim that these findings will apply to schools in other communities. The purpose of this study was to discover whether or not group differences existed from a psychological point of view. It is apparent from the above results that such differences do exist and that some of these may be detected by the use of a group intelligence test. With such differences studied and measured, school authorities will find them of value in the educational and vocational guidance of the students who naively choose to take special High School courses.

THE TESTIMONIAL AS AN ADVERTISING APPEAL

By EGBERT M. TURNER, College of the City of New York

The purpose of this paper is to attempt to determine the value of the testimonial method of advertising from the viewpoint of the permanency of the testimonial writers' interest and belief in the worth of the article being advertised. The paper does not attempt to analyze the number of sales made by such appeals, but attempts to determine whether or not the writer of the testimonial:

1. Continues to believe in the worth of the article he first recommended.
2. Whether he bothers to answer the letters of those requesting advice.
3. Whether or not men and women differ with respect to the ways in which they as testimonial writers react toward letters written to them for advice.
4. Whether or not people with the medical degree differ from an unselected group in the way in which they as testimonial writers for an article of a medical nature react toward letters written to them for advice.

A business concern may, without decreasing its sales, spend thousands of dollars yearly for electrical signs which amuse and interest spectators, who often are unable to tell the name of the article being advertised a few minutes after enjoying the dazzling lights. However, a concern depending mainly upon the testimonial method of appeal cannot afford to invest money in printing and circulating these testimonials unless it is certain that these appeals are not actually interfering with and decreasing its possible sales.

Many advertising appeals do not lead to an immediate response, but require deliberation, comparison and choice. Professor H. L. Hollingworth (Advertising and Selling), calls such advertising appeals which bring into activity the higher thought processes, such as reasoning and decision, the "long circuit appeal." He calls those advertisements which lead to almost immediate response, because of their appeal to some fundamental interest, feeling or instinctive tendency, the "short circuit appeal."

Analysis of the method of advertising the article studied in this paper indicates that the appeal is a combination of the

short and long circuit appeals. The article being of a medical nature immediately appeals to those suffering the ailments it claims to cure. This short circuit appeal thus causes the observer of the testimonial pamphlet cover to read its contents. From this point on the short circuit appeal is overshadowed by the long circuit appeal, i. e., argument, reasoning and decision. If decision is not made, either to buy or not to buy, the reader often writes to one of the testimonial writers for advice,—thus protracting the appeal. Should the answer fail to come, or be unfavorable toward the article, the appeal is probably lost, and from the viewpoint of the concern a sale is lost,—hence the need of investigating the worth of the testimonial as an advertising appeal.

The business house for whom this study was made has 279 testimonials in eight different pamphlets. These pamphlets are sent to people whose names are secured in the usual mail order house method, and also to those inquiring directly about the article. Of the 279 testimonials 53 were written before the year 1916, 85 in the year 1916, 97 in the year 1917 and 44 during and since the year 1918.

In making this study the same type of letter was used in approaching the testimonial writers, stating that the writer had noted that he or she was an enthusiastic user of the article advertised and would appreciate their advice about it as he was contemplating the buying of it. A stamped envelope was enclosed in all of the letters, and "Mr." signed on all letters addressed to men and "Miss" on all letters to women. The reason for this precaution was that a sensitive person might object to write to one of the opposite sex about a medical article.

The replies and returned letters were classified as:

1. Good,—namely, all those replies recommending the article, and which might influence one to buy it.
2. Poor,—those which would cause one not to buy the article.
3. Returned,—all letters returned by Post Office as unable to locate.
4. No answer,—all letters which brought no response and were not returned.
5. Deceased,—letters returned advising that person addressed had deceased.

The limitations of this classification are evident. The replies classified as "Good" ranged in their advice all the way from a mere approval of the article to an enthusiastic "all

power" belief in it. Likewise, the letters classed as "Poor" ranged from a mere disapproval to a strong belief that the article was detrimental and dangerous to health. The letters were judged by four people,—the Medical Advisor to the Institution, the President, the General Manager and the Manager of the Mail Order Department. For the most part, their opinions were unanimous. Undoubtedly a ranking on the basis of 1 to 5 would have been more desirable, but it is extremely difficult to get business people to use and think in fine measures of gradation. Naturally those letters classified under "Good" were the only ones which could be considered as making a positive appeal, while those classified as "Returned," "No answer" and "Deceased" certainly did nothing to encourage the inquirer to buy, while those classified as "Poor" would tend to cause the inquirer not to buy.

The following table gives a statistical analysis of the returns for the 279 letters written:

TABLE 1
SUMMARY OF REPLIES AND RETURNED LETTERS.

	Men, Women & Institutions		Men		Women		Institutions		Men, Women and Institutions Before 1916 Year 1916 Year 1917 During & after 1918							
	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%
Good.....	138	49.4	90	47.9	46	56.1	2	22.2	17	32.1	39	45.9	56	57.7	26	59.1
Poor.....	19	6.7	9	4.8	8	9.8	2	22.2	6	11.3	5	5.9	6	6.2	2	4.5
Returned..	52	18.6	40	21.3	11	13.4	1	11.2	13	24.5	21	24.7	11	11.3	7	15.9
No answer.	55	19.7	35	18.6	16	19.5	4	44.4	13	24.5	15	17.6	19	19.6	8	18.2
Deceased.	15	5.6	14	7.4	1	1.2			4	7.6	5	5.9	5	5.2	1	2.3
Total.....	279		188		82		9		53		85		97		44	

It will be noted that only 49.4% of the replies were favorable; that is, about one-half of the letters written to the original testimonial writers elicited favorable replies. Of the replies, 6.7% were of the nature which would discourage the potential buyer. The remaining 43.9% of letters failed to bring forth a response and thus made no selling appeal whatever. If anything like these percentages hold good in other concerns the testimonial method of appeal is very wasteful since it favorably impresses only about one-half of the prospective buyers who have actually shown an interest in the article to the extent of writing for advice about it.

Furthermore, it will be noted that the more recently written testimonials are more effective than the older ones. Of those writing testimonials before the year 1916 only 32.1% replied favorably; of those writing during the year 1916, 45.9% replied favorably; of those writing during the year 1917, 57.7% replied favorably; during and since 1918, 59.1% replied favorably. The favorable responses for the different periods range from 32.1% to 59.1%. This calls attention to the imperative need of continually revising testimonial pamphlets, but even the highest percent of 59.1 seems to the writer to be amazingly low.

To speak with certainty of this apparent wastefulness it would be necessary to compare the sales resulting from the testimonial appeal with the sales resulting from an appeal free of testimonials. This study might be conducted by sending two unselected groups two different forms of appeal. An analysis of the number of sales would then indicate the relative values of these two methods, and also whether or not the fact that from 40 to 68% of testimonial writers fail to respond to inquiries about the article they have recommended decreases the sales below the number made by the other appeal. It is believed by some that people prefer to write testimonial writers instead of the business concern, feeling that the advice will not be biased, and also that they escape the many follow-up letters sent by the usual mail order houses. Should the analysis indicate the value of the testimonial method it would then call attention to the need of devising a way of increasing the percentage of favorable replies.

It is of value to determine whether or not there are any characteristic differences between men and women with respect to the ways in which they behave as writers of testimonials. Of the 279 testimonials 188, or 67.4% were written by men; 82, or 29.4%, by women and 9, or 3.2%, by representatives of institutions. The concern claims that there was no conscious selection of men's testimonials, but that when the pamphlets were compiled the good testimonials were selected regardless of sex. It may be that women are somewhat more reserved than men about recommending an article of a medical nature. While the concern has no record of its sales according to sex still the general belief is that more men than women buy it.

It will be further noted that 47.9% of the men and 56.1% of the women answered the letter, requesting advice, favorably. That is, 8.2% more women answered the letter favorably than men. On the other hand, it will be noted that 4.8

of the men wrote unfavorable replies and 9.8 of the women wrote unfavorable replies. That is, 4% more women wrote unfavorable replies than men. While these differences are small still the indications are that a greater percent of women wrote both favorable and unfavorable replies than men.

Among the testimonial writers were a few doctors, but the number of men and women holding the medical degree is too small to permit any valuable generalization, still the table below shows indications that are worth considering:

TABLE 2
COMPARISON OF REPLIES FROM DOCTORS AND NON PROFESSIONALS

	Doctors	Non-Professionals		
Good.....	8	44.4	130	49.8
Poor.....	4	22.2	15	5.8
Returned.....	4	22.2	48	18.3
No Answer.....	2	11.2	53	20.3
Decreased.....			15	5.8
Total.....	8		261	

It will be noted that of the 18 doctors who first recommended the article only 44.4% wrote favorable replies to those requesting their advice, as compared with 49.8% of 261 people other than doctors, who answered favorably. Also that 22.2% of the unfavorable replies were made by doctors as opposed to 5.8% of the unfavorable replies written by others than doctors. This seems to indicate that from the viewpoint of a business concern those of the medical profession are less dependable as testimonial writers regarding articles of a medical nature.

CONCLUSIONS

1. Approximately an average of only about 50% of testimonial writers answer favorably in response to inquiries from prospective buyers of the article in question.
2. About 6 or 7% of testimonial writers give unfavorable and unsatisfactory replies in regard to an article they at first recommended.
3. Recent writers of testimonials are inclined to write more

favorable replies,—hence the need of frequently revising testimonial pamphlets.

4. The difference between men and women with respect to the ways in which they behave as writers of testimonials is small, but the indications are that women write a greater percentage of both favorable and unfavorable replies with reference to testimonials.

5. From viewpoint of a business concern, those of the medical profession are less dependable as testimonial writers relative to an article of a medical nature than unselected people.

EFFECT OF LETTERS AND SYLLABLES IN PUBLICITY

By EDWARD S. JONES, Oberlin College

It is patent from a casual notice of trade names and other symbols used in advertising that certain forms are much more desirable than others. Comment regarding the characteristics of letter and syllable forms in language go back at least to the rhetoric of Dionysius of Hallicarnassus who undertook to analyze the affective quality of various Greek letters. Unfortunately a study of this sort for one language does not carry over into other languages well, as many letters are pronounced somewhat differently in different tongues which may have derived their forms from the same roots.

Some one has suggested that the affective quality of articulated forms is closely related to, if not contingent upon, the affective quality of predominant tastes in that part of the mouth effecting the articulation. The harsher forms, for example, such as "g" involve the back of the tongue and consequently are connected with the sensation of bitterness assuming an unpleasant flavor. It is presumed that early in the language experience of the child, when the taste sensations are in relatively pure form, these affective conditions are started. Words denoting disgust, and the emotion of fear (including certain religious utterances) are cited as evidence.

Of course it is possible that other influences are much more potent in such affective tendencies (which are really conditioned reflexes) than the sensations of taste. This is specially probable in consideration of the recent agreement on the part of investigators that taste "localization" on the tongue is not nearly so definite as it was once thought to be. Letters involving extensive movements of the uvula or throat, such as "k," "g" and "q" are utilizing the same muscles which are active in sudden emotion, particularly despair. The sonants "b," "d," "g" and "v," which involve a vibrating larynx as well as throat movements may be specially correlated with sustained excitement. The "l," "r," and "s" forms seem to include front of mouth movements, and probably a tendency to widen the mouth, and may thus be correlated with the emotions of love,—hence engender pleasure. The

"s" form, however, is considered quite unpleasant by Dionysius on account of the accompanying hissing sound, which should remind us that auditory imagery must of course complicate any effort to correlate feeling quality with the articulation peculiar to language forms.

But this is all speculative. If a connection between our simple letter forms or syllables and affective states does exist, it should be subject to investigation. No doubt many authors have an unusual sensitivity for the "feeling" of forms in prose or poetry.¹ It is possible also that experienced advertisers are able to select symbols which are calculated to be not only pleasant but also easily remembered as compared to other forms. Such differences in *affective* and *memory* "value" in letters or syllables should be experimentally determined and objectively recorded to some extent at least, if they are worthy of our attention in matters of publicity.

The experiment to be recorded is a mere outline of what is experimentally feasible. Only 15 subjects were used, and for purposes of simplifying data, and to study special types of articulation, we confined the experiment to certain letters only, neglecting the vowel "i" for example. We used only nonsense syllables,—so far as possible those which had no sound or sight similarity to real situations. This type of stimulus has two advantages: it minimizes the significance of associations and early affective conditioned reflexes arising therefrom; and it also tends to approach a practical condition in advertising. Many trade-names are nonsense syllables or combinations of such syllables so far as the public is concerned. It is one way of making a symbol vivid and distinct (e. g. rit, tirro, hag, zet, loju, kef, pax, raz, vital, zim).

Our method was to make up a list of syllables, 50 in all, and present them in pairs to the subjects. Each syllable occurred twice, each time with a different mate, once the first of a pair and the other time the last. This presumed to eliminate differences due to position, and to compensate somewhat for the influences of specific syllables or sounds on others. A subject was asked to force a distinction between the two syllables shown, and to grade each differently on a scale of 3 points. That is, it was the paired associates method, allowing at the same time for some grading, so that if the two syllables were markedly different in affective quality, one might be graded "3" and the other "1." Each subject was asked to think of "3" as significantly pleasing and well liked; of "2" as above average in affective tone, at least not dis-

¹ Cf. Bliss Perry, "A Study of Poetry," 1920.

pleasing; and "1" as below average, or positively distasteful. A more extensive scale of measurement, say a five place system, might have been used to advantage, though it would require a longer time for decisions which might in turn have allowed for the development of further associations in the case of syllables used. After each judgment had been made, the subject was asked not only to indicate the number to be assigned to the right and left syllable, but also indicate after a number a plus sign if there was a pleasant association or meaning attached, and a minus sign if there was an unpleasant association. There were only three out of the fifty syllables used which seemed to call up associations for the majority of the group; although of course this method does not account for unconscious affective influences arising by the conditioned reflex route.

The second part of the experiment came ten days later, without the subjects' knowledge that they would be expected to remember the syllables they had seen. Each of the syllables had matched with it a confusing syllable, only one of the three original letters being altered. The one hundred syllables were presented to the subjects in random order, and the subjects were asked to indicate those of the group which they had seen before. They made a distinction also between those they were sure of and those they thought they had seen, but were less sure about.

The results are suggestive. Out of the 50 original syllables shown, the following—in the order from high to lower affective value—turned out to be distinctly pleasing to the subjects:

"Nep, lon, dur, paz, des, zef, tor, dos, reg, ros, ber, tur, tul, ket, tov,"

Those syllables most unfavorably considered for the 15 subjects, in order from medium to the least pleasurable forms were:

"Bod, seg, sev, kuv, tas, kaz, ruv, sab, rav, geb, vab, vud, ged, gur, gak."

On account of the large number of "neutral" syllables, of little difference in affective value, in between these lists, there is unquestionably more difference between "Nep" and "tov," or between "bod" and "gak," than between "tov" and "bod." The syllables in the medium affective range, in the order of more pleasing to less pleasing were:

"Dav, lof, teg, keb, bep, gof, sef, teb, lug, ret, mul, pev, sud, zud, reb, dap, kop, bov, sof, kut,"

An inspection of those syllables accompanied by associations leads one to feel that the matter of favorable or unfav-

orable associations was not very significant in fixing the order in the series of any syllable. For example, the form "dur" had 4 positive (favorable) associations out of the 30 chances (2 for each subject), and 2 negative associations. The form "lug" had 2 positive and 6 negative associations attached; while "sud" with the largest number of associations noted had 10 negative and 6 positive marks. The form "gof" which stands at about the middle of the series had 8 positive and no negative associations recorded against it, while the form "gak" which is far below all the others in its affective quality only had 7 negative associations noted. All the other forms had 5 or fewer positive and negative marks, out of a possible 30, indicating associations.

The following preliminary conclusions seem to me justifiable: Forms involving the letters "t," "d," "l" and "r" tend to be pleasurable. There is not enough evidence for "n" and many other forms. Forms involving "g" and "v," and to a less extent "k" and "s," tend to be unpleasurable. It is interesting that of the two labial occlusives, "p" and "b," the latter seems to stand less favorably, perhaps because it involves more palate movement.

The beginnings of such simple forms seem to be more indicative of the probable affective tone than the endings, or the vowel; i. e. there is more regularity between the presence of certain letters and the affective quality of the syllables as a whole, when we consider first letters.

A very important question that arises is:—do pleasing letter and syllable forms when put in combination always remain pleasing, and will the insertion of an unpleasant form necessarily tend to lessen the affective tone of a combination? Only a few random cases of the combinations of such syllables were tried; i. e. the presenting of a five or six letter form made up from these syllables. On the basis of the results obtained from a few subjects, it appears that the affective quality of a combination of two syllables can in general be predicted from the affective character of the components, but that the first of the two combinations is far more potent in determining the type of affective tone than the second. The forms "torav" and "desev," for example, ranked fairly high, though the final syllables of each were considered quite unfavorable,—as noted in the above lists. It is also possible that many combinations, whose components are separately disagreeable, may be of average affective value at least. This seemed to be true of the form "ravab."

The relationship between letters and syllables and *memory* value is not as close as in the case of the affective quality. As already indicated, the memory value of a syllable was gauged by the number of persons recognizing it, among other syllables specially intended to confuse, after a ten day interval of time. Syllables of highest memory value, beginning with the most easily recognized stand in this order:

"Bep, sud, sof, tov, reb, tul, ruv, sev, tas, bov, des, kuv, pev, paz, sef,"

Those standing lowest in memory value in the order from medium to the lowest are:

"Gof, lon, kut, dos, bod, teg, ros, mul, ket, kaz, keb, geb, ged, dav, vab."

Associations are of slightly more value in the case of recognition than in determining affective tone, the 6 women apparently benefitting more than the 9 men subjects in this respect. Seven out of 15 syllables standing highest in memory value were designated as having associations by one or more subjects, whereas only 4 of the 15 syllables lowest in memory value had associations.

It seems probable that syllables starting with "s" and "t" are favored as to memory value. This seems also true of "r" and "p" sounds. The endings in "v" are apparently specially high in memory value. It is interesting that all of these letters are distinctly 'front of mouth' forms, involving lips or tongue tip, though "v" is also a palate form. The letters "g" and "k," standing low in memory value, are distinctly 'back of mouth' forms. It is apparently not the amount of articulation in a form which determines its memory value as some authors have suggested.

There is perhaps also some connection between vowels and memory value. Of the seven syllables most readily recognized only two have "a" or "e" as vowels, whereas the lowest seven syllables all have "a" or "e" centers. However, except for the extremes, this relationship does not hold and it is therefore of doubtful importance.

Apparently there is little or no relationship between pleasure value and memory value in simple syllables. The correlation between these two factors measured objectively, as described above for the 50 syllables, is barely positive—less than .1. Those syllables with the highest memory value are apt to be mediocre or neutral in affective quality. Several syllables which were easily confused in later recognition were quite agreeable, e. g. "lon, dos, ros, ket, and dav." However there

is one serious source of error in considering the memory value of single syllables which should not be minimized; i. e. the difficulty of testing rigidly and comparably each of the 50 syllables used. It was obviously impossible to incorporate into the recognition test, of 100 syllables, 50 confusion syllables each equally different from one of the older 50 syllables which had been judged for affective value.

One additional point of interest was studied,—the relationship between the pleasure rating, and also the memory value, of the initial 50 syllables and the commonness of these syllables in the ordinary word forms of every day life. An abridged dictionary was referred to, and a notation made of the relative frequency of each of the syllables standing high and those standing low in affective value, and in memory value. There was no relation evident between the prevalence of syllables in our every day language and the memory value of these syllables. Of the ten nonsense syllables lowest in memory value, 4 were new or unfamiliar and 4 others were only slightly used; of the 10 syllables highest in memory value 3 were new and 4 only slightly used.

Probably more of a relation exists between the affective character of syllables and their frequency in language forms. Especially is it true that totally unfamiliar syllables tend to be disagreeable. Six out of 8 of the syllables lowest in memory value were practically unused syllables. It is quite possible that the fact that these syllables are unfamiliar does not cause their unpleasantness; on the contrary language usage may have "selected" for its permanent forms those syllables which were the most pleasing. Obviously, however, any disagreeable form repeated often enough will lose much of its unpleasantness.

In summary, it is evident that the affective and the memory value of different letter and syllable forms can be experimentally studied to the advantage of Applied Psychology. The results obtained from a preliminary experiment are in line with opinions occasionally expressed. Forms involving the liquids "l" and "r" tend to be pleasing. The dentals "t" and "d" are distinctly preferable to the harsher occlusives "g" and "k," and to the fricatives "v" and "s." Initial letters, or syllables, of a combination are specially significant in determining the affective status of that combination. Forms which necessitate lip and 'front of mouth' movements have the advantage in memory value. There is no simple relation between the pleasure value and the memory value of different

forms; but on the average the 'front of mouth' forms seem to have the advantage in both respects. The relative familiarity which we have with syllables in a language does not seem to affect their memory value, nor their affective tone to a significant extent; however, the most disagreeable syllables are apt to be unfamiliar.

Such a preliminary experiment can but point to future lines of investigation, detailed and objective. Some of the questions for possible research are as follows:

Are the above results confirmed with more extensive experimentation with many more combinations of letters and syllables, and when different methods are used to measure the affective and memory values?

What is the influence of repetition on the affective tone of syllables? May certain relatively neutral forms become quite pleasing?

Will the affective tone of a syllable, and possibly its memory value, differ markedly when we consider different articles for which it may be a symbol (e. g. automobiles vs. food stuffs)?

Are there emotional experiences other than simple affections which should be taken account of in favoring a letter or syllable form in considering publicity? i. e. may certain forms be more "solid," or more "elegant," than others regardless of their pleasure value?

SUBSEQUENT HISTORY OF E—; FIVE YEARS AFTER THE INITIAL REPORT

By LETA S. HOLLINGWORTH, CHARLOTTE G. GARRISON, and AGNES BURKE,
Teachers College, Columbia University

IN THE JOURNAL OF APPLIED PSYCHOLOGY for 1917 a report was made of the mental status and educational achievement of a boy, E—, at that time 8 years old.¹ Reference to this initial report will show that at the time of first mental examination E— showed an IQ of at least² 187 (Stanford-Binet), and that he was then doing superior work in the 6th grade of the Horace Mann School. In this school the median IQ of the pupils is about 116, and the median age for pupils in the 6th grade there is 11 to 12 years.

It is the purpose of this paper to give a brief account of the development of E— five years after the initial report. Mental and physical measurements have again been made, and scholastic achievements have been ascertained, as follows. Additional facts, also, concerning E—'s genealogy are herewith presented. It is a question of good taste, and perhaps to some extent of ethics, as to whether an investigator may give the name of a living subject of scientific research. Since E— has already become well known in the community where he resides, and since the facts as now set forth are undoubtedly sufficient to identify him, there seems no real point in omitting his name from this report. Permission has therefore been sought and granted for the use of E—'s name, which is Edward Rochie Hardy, Jr.

Additional data on Genealogy. In the initial report data were given concerning the parents of Edward. Fairly complete genealogical records of maternal ancestry are avail-

¹ Garrison, Burke, and Hollingworth, *The Psychology of a Prodigious Child. Journal of Applied Psychology*, 1917.

² "At least" because E— on first examination passed some tests at the highest level afforded on the Stanford-Binet. Had still more difficult tasks been provided, he might have succeeded with some of them, though in the opinion of the examiner he probably would not have done so.

able.^{3 4 5}. Five persons bearing the surname of the mother settled in New England before 1650. These were probably all related to each other. The individual who was Edward's direct ancestor first appeared in New England in 1639, and settled in Cambridge, Massachusetts. This family attained great distinction in the six generations recorded in the New England genealogy. A son of the first ancestor in America was a royal councillor, and the greatest merchant of his day in New England. A grandson was royal governor of Massachusetts, and later of New Jersey. He was also a patron of learning. A great-grandson was chief justice and lieutenant governor of Nova Scotia. A great-great-grandson was a royal councillor of Nova Scotia, some of whose children settled again in England, of whom a son was a distinguished naval officer, attaining the rank of rear-admiral in the British navy.

These genealogical records, and other records of New England families which intermarried with this family, have not been brought up to date. The maternal surname appears first about 1176 in the records of England, and was apparently Norman-French in origin.

Mental Measurements. In the spring of 1920 Edward took the Thorndike Mental Tests for Freshmen, for entrance to Columbia College. An official letter from the Director of Admissions at Columbia College states that, "In the Freshmen Tests he (Edward) was number two, out of 483 entering Columbia College." Edward was then 12 years 0 months old, the median age of his competitors being about 18 years.

On September 29, 1921, Edward was examined mentally by Dr. Hollingworth, by means of Army Alpha, Form 5 and Form 6, for purposes of recording his mental development. On Alpha, Form 5, he made a score of 194 points, finishing several of the tests before the time limit, without error. On Form 6 of Alpha, which was taken subsequently on the same afternoon, his score was 201 points; and here too some of the tests were finished in about two-thirds of the time allowed, without error. As the method of scoring Alpha does not provide for a time bonus, this cannot be taken into account in the formal score.

³ Bartlett, J. G., *The Belcher Families in New England*. N. E. Hist. and Gen. Reg., vol. 60, pp. 125-136, 243-256, 358-364.

⁴ Belcher, Jonathan, *The Belcher Papers*. Coll. of the Mass. Hist. Soc., vol. 6 and vol. 7.

⁵ Appleton's *Cyclopædia of American Biography*, vol. I, 1887. (Jonathan Belcher.)

Physical Measurements. On September 29, 1921, measurements were taken of height, sitting height, and weight, as follows: ht.=64.2 inches, sitting ht.=31.7 inches, wt.=166 lbs., (with summer clothing). Health remains most excellent. This child has never had a serious illness of any kind.

Scholastic Achievement. In the spring of 1917, Edward finished the 6th, 7th, 8th and 9th grade work at the Horace Mann School. He was then just 9 years old. Thereafter he attended The Friend's Seminary, and graduated from the high school there in the spring of 1920, with an excellent record, and excess credits. He was then 11 years 10 months old.

By this time, also, Edward had passed the comprehensive examinations of the College Entrance Board for Harvard College, one of the present writers having seen the official communication from Harvard authorities, making statement to this effect. Edward's maternal ancestors had attended Harvard (one of them having graduated from there at the age of 18 years, according to the records), but Edward expressed a desire to attend Columbia, and received permission to take the mental tests with the applicants of 1920. He was admitted to Columbia College with the freshmen of 1920, with 14 points of advance credit toward a B.A. degree.

The present writers have at hand an official statement of this child's academic status on June 1, 1921, at the age of not quite 13 years. He had then 46 points of academic credit toward a B.A. degree in Columbia College. During his freshman year he made 32 points, maintaining consistently a grade of B, except in two subjects. In *physical education* his rating was C, and in *contemporary civilization* he made A the first semester. (It is worth saying that in *contemporary civilization* the final examination⁶ had been objectively standardized by Mr. Ben D. Wood, expert in educational measurement, and did not depend on the estimate of instructors). These marks have the usual significance: A = excellent, B = good, C = fair, D = poor, F = failed.

Edward attended the Summer Session of 1921 at Columbia, making five credit points, all A grade, and is at this date, September, 1921, a sophomore in Columbia, with many points of credit in advance of minimum sophomore status. In addi-

⁶ "Hardy's score on the objective examinations of both terms was in the highest tenth of the highest percentile." B. D. W. (The instructor's estimate, as well as the result of the objective examination, enters into the *term* grades recorded, and indicated above, in this course; so that the final grade in the second semester is but B.)

tion to having passed the comprehensive examinations for college entrance, he has passed the examinations in trigonometry, solid geometry, chemistry and physics. He is 13 years 3 months of age.

Extra-Curricular Activities. Edward was, of course, a conspicuous freshman, because of his extreme youth, and he was "hazed" by the sophomores for refusing to wear a prescribed necktie. One of the New York newspapers commented on his conduct under hazing as follows:

"He has demonstrated that he is, nevertheless, a regular fellow. He did it first by bringing about a conflict, in which he himself was the much-buffed prize of battle, and then by glorying in his bruises instead of making them the basis of a grievance.

"He is a good sport, as well as a good scholar, and being both he ought to go far."

Edward also participated in the class play, given in 1921, humorously consenting to impersonate himself.

Manual work has no more charm for Edward than it had when he was 8 years old. That he *can* work with his hands and with materials when motivated is suggested by an incident connected with the liberty bond drive. His teacher relates that Edward wanted to pay for his own bond; so he made jelly, working at it till very good jelly was made, and sold it for the purpose specified.

In recent years there has developed a keen interest in detective stories.

Comments of Teachers. Comments from Edward's teachers during the last five years are indicative of their estimates: "The regular course of study has been so easy that he has, in several subjects, notably English and History, accomplished a great amount of voluntary work outside the course." "An excellent mixer with older students." "His weekly visits have been a pleasure and anticipation, and his ability to understand without English the spoken Latin and the authors as I have read them aloud to him, has been extraordinary." "Has done very remarkable work in science, particularly in theory." "I predict for him a great scholastic record in college." "I consider it a privilege to have had something to do with teaching him." "Possesses a *power* in Latin that few persons ten years his senior can boast." "Has shown devotion to the best interests of the school."

Summary. In the five years which have elapsed since Edward was first tested mentally, he has shown no tendency

to become mediocre. His gifts have not grown less. He maintains his superior status in mental tests. As for achievement, he has passed during this interval from the 6th grade of the elementary school, half through the second year of college. Average children, the country over, born when this child was born, and measuring 100 IQ when he measured 187 IQ, are now in the 7th grade of the elementary schools. Edward's development in the traits and achievements measured may be summarized as follows:

Born June 17, 1908. (According to official birth certificate, as seen by Dr. Hollingworth, and according to statement of mother, as received by present writers.)

	E.	*Norm
<i>November 4, 1916.</i>		
Height.....	54.3 inches.....	49.5 inches
Weight.....	89.3 lbs.....	54.2 lbs.
Intellectual.....	IQ 187.....	IQ 100
status	(Stanford-Binet)	(Stanford-Binet)
School.....	6th grade.....	*3rd grade
status	elementary school	elementary school

<i>September 29, 1921.</i>		
Height.....	64.2 inches.....	58.2 inches
Weight.....	166 lbs.....	89.5 lbs.
	(light clothing)	
Intellectual.....	194 points.....	47 points
status	(Alpha, Form 5)	(Alpha)
School.....	4th semester.....	*8th grade
status	college	elementary school

It is an interesting theoretical question as to how far human intelligence may vary from the norm, in the direction of superiority. The case of this child has been placed on record largely because it seems probable that such cases represent very nearly the extreme possible limit of variation in the human species, as it now exists. At 8 years of age his IQ stood at +11 P.E., (1 P.E. being ± 8 IQ, according to Terman's distribution). The probabilities are usually regarded as slight that cases beyond 5 P.E. will occur. Perhaps the range in human intellect is much greater than probabilities would lead us to guess.

Edward still wishes to be a minister, and go abroad as a missionary. To this end he interests himself especially in history, the languages, and anthropology.

* Note: Norms for height, weight and school status are taken from Bird T. Baldwin, as established on children from Horace Mann School, Francis Parker School, and the Elementary and High Schools of The University of Chicago, where the children are above average (for the population at large) in intelligence.

Since the initial report of this child's qualities, readers have occasionally asked with what meaning the word "prodigious" was used in reference to him. It was used in the dictionary sense of "wonderful," "extraordinary."

In these reports there is no intention to approve or to disapprove the educational regimen pursued. Who knows what should be the educational treatment of a child standing at +11 P.E. in intellect? The sole intention is to record facts concerning the identification and development of a deviation so extreme that the chances are theoretically almost nil that it would occur at all.

THE SIGNIFICANCE OF UNWILLINGNESS TO BE TESTED

By F. B. KNIGHT, State University of Iowa

As a pedagogical device in introducing the 'mental test' to the teachers of a county in Iowa, I had the teachers take Thorndike's College Entrance Examination, Part I. At the time of taking the test the teachers were given the choice of signing their names or not as each saw fit. They were assured that no administrative use of any kind would be made of the scores. The superintendent would not be informed of the relative success of any teacher. However, those teachers who signed their names to their tests would be told what scores they made and how their scores compared with scores made by college students, stenographers, soldiers, and the like.

There were 65 teachers in the group, 39 signed their names, 26 did not. The distributions of these two groups are as follows:

Score	Frequency of scores on signed papers	Frequency of scores on unsigned papers
35		1
40		
45		2
50	1	
55	1	3
60		1
65		2
70	2	2
75	1	3
80	3	4
85	4	
90	1	2
95	5	2
100	3	5
105	4	1
110	5	
115	2	1
120	3	
125		
130	1	
135	2	
140	1	

The average score of the signed papers was 100.

The average score of the unsigned papers was 75. Only

212 SIGNIFICANCE OF UNWILLINGNESS TO BE TESTED

four of the 26 unsigned papers equalled or exceeded the average of the signed group. Only 5 of the 39 signed papers were as low or lower than the average of the unsigned tests. We have very definite contrast in performance between these two groups. Chance variations even with as small a group as this would be an unlikely explanation of the difference of the two distributions.

Although we have too few cases upon which to base a certain conclusion it is evident that unwillingness to sign one's name to a mental test may be an unconscious confession of probable inability to do well. Willingness or even desire to know how you compare with others is associated with relatively good scores. This explanation seems to me more reasonable than that unwillingness meant becoming modesty or willingness inordinate forwardness. Should not one's attitude toward a test be of definite diagnostic aid?

NOTES AND COMMENTS

Announcement of the Psychological Corporation is made in a recent number of *Science*.

In answer to the letter addressed last year to psychologists in regard to the plans for the corporation, 161 replies were received. These were all favorable with the exception of two letters from psychiatrists. There was general expression of appreciation of the importance of an organization that will maintain adequate standards in the applications of psychology and will use any profits there may be for psychological research.

The directors have authorized the establishment of branches wherever a director or three members of the American Psychological Association are ready to take an active interest in the work. We shall be glad to receive word from those to whom this memorandum is addressed in regard to the formation of a branch in their neighborhood.

The stock of the corporation is to be held by psychologists active in its work. The directors have authorized the sale of from one to three shares to competent psychologists. The stock will be sold for \$10 a share and may be taken over by The Psychological Association at \$100. The principal capital of the corporation is the ability and interest of those who take part in its work, and a psychologist who subscribes \$10 for a share is expected to give time worth \$90 before disposing of the stock, though only as a general understanding, the decision regarding which is left entirely to the holder of the stock. Those to whom the memorandum is addressed who want one share of the stock may send a check for \$10 to the treasurer of the Psychological Corporation, and the stock will be forwarded. If they want one or two further shares application should be made and the shares will be allotted if available.

J. McKEEN CATTELL, *President*.

The Psychological Corporation,
2617 Grand Central Terminal,
New York City.

February 20, 1922.

Information Course on Science of Management and Applied Psychology.

From August 21 to August 26, 1922, an Information Course on the Science of Management and Applied Psychology will take place at the "Orga-Institut"—an Institution devoted to the Science of Management and Applied Psychology in Berlin.

This Course will include lectures in the lecture-hall of the Institution and practical exercises in the laboratory as well as inspection trips with descriptive lectures.

It is planned to visit the following Institutions: the Psychotechnical Laboratory of the United Berlin Street Car Companies (employment psychology and vestibule school), the Psychotechnical Laboratory of the Head Postmastery, the Laboratory for Industrial Psychology at the Technical High-School-Charlottenburg, the Psychotechnical Lab-

oratory of the Berlin Office of Labor, a number of Psychological laboratories of large Berlin industrial concerns, the Institute of Applied Psychology directed by Geheimen Sanitätsrat Dr. Moll (specially for pathological cases) as well as some other places if the time set suffices.

The languages spoken at the Institute are German, English and French.

Only a limited number of students will be admitted, in order to give every member of the course the utmost possible consideration and to enable him to get full insight into the problems to be dealt with.

As all applications will be considered according to the date of receipt, it is advisable to give notice of attendance at the earliest possible moment.

The fees are for:

Dutch, 75 Gulden

Swiss, 150 Francs

Danes, 150 Crowns

Swedes, 150 Crowns

Norwegians, 125 Crowns

Englishmen, 5 £

Americans, 25 Dollars

Italians, 400 Lire

Spain, 150 Pesetas

For other Nations fees will be in accordance with these rates.

BOOK REVIEWS

HARRY DEXTER KITSON. *The Mind of the Buyer, A Psychology of Selling*. New York, The Macmillan Company, 1921. Pp. x+211. Price, \$1.50.

Not to treat of the technique of selling, but of certain psychological principles related to the mind of the buyer, such is the purpose of this book. Most of the illustrative material, so far as it bears directly on selling, has to do with advertising. The outline treats of the usual stages in the sale, with good analysis of the important topics: attention, interest, desire, confidence, decision and action. In addition, a final chapter on the stage of satisfaction brings in an emphasis which is one of the best things in the volume. The salesman, advertiser, and sales correspondent, for whom the book is written, will find it quite readable and brief. The author has a facility for catchy phrases which stands him in good stead. Occasionally the psychological terminology clings to the pages in spite of the author's attempt to escape it. Dwelling on images, for example, seems unnecessary.

The psychologist will be interested in Kitson's excellent use of the historical method in studying advertising in order to discover the drift of progress through the decades. That advertising has made progress by its hit or miss methods is well shown by his studies of the decrease in the use of superlatives, decrease in the use of all capitals for headlines, and analysis of the kind of information used in advertisements. One regrets that more concrete examples of his principles from the field of personal selling were not available to the author. The illustration of the method used by a canvasser to sell a cooking product to a housewife is good, but is somewhat lonesome. Since the book meets a popular demand, often uses figurative language admirably for this purpose, and is generally sound in its principles, the reviewer may perhaps be most useful in wishing that somebody would build a different kind of book based on a psychological analysis of a collection of methods used by successful salesmen and illustrated with details of the different methods.

J. B. MINER.

PATON. *Human Behavior, in Its Relation to the Study of Educational, Social and Ethical Problems*. New York, Charles Scribner's Sons, 1921, pp. 465. Price, \$6.00.

To the student of applied psychology no work would be more welcome than one on human behavior which should open a broad highway into that complex and difficult subject. Such a highway is, however, too much to expect at the present time. The practical control of human behavior has been so long a major task of so large a part of mankind that all the easy portions of the field have long been familiar, while its deeper penetration has had to wait for the advance of knowledge in other fields, especially in those of biology and physiology, if indeed the means are even yet available for any great advance of a rigidly scientific sort. For this reason it is enough if Dr. Paton blazes us a trail; in fact, he himself refers to his work as an "introduction," and in his exposition finds himself often brought to a stand by the absence of necessary information not yet available.

The main thesis of the book is that the key to the complexities of adult human behavior is to be found in the simpler sorts of behavior appearing in the lower, and even lowest, animals, in the earliest stages of human life (including the prenatal) and in the disordered processes of mental and nervous disease. The study of behavior involves the study of the responses of the total organism to the total environment. One must go not only back of the brain and spinal cord to the autonomic nervous system, but back of that to the endocrine glands, the muscles, the vegetative system and even to the most rudimentary adaptations of living matter; and above all one must not separate the mental from the physical.

The first of Dr. Paton's fourteen chapters traces the historical development of this biological-physiological point of view. The next eight discuss behavior (and personality, its unified expression) in its dependence upon the bodily mechanisms of adjustment and control, upon "trends" in activity and upon habit-formation. The next two deal with the typical problems of degeneration in old age and of the imperfect organization of personality appearing in conflicts and dissociations; chapter twelve with the method to be used in studying personality; chapter thirteen with education; and chapter fourteen with the more general sociological and international results that may be expected to follow from the study of the causes of human action.

To the general reader the chapter on Education will perhaps prove most interesting and can hardly fail to win the hearty approval of any who know what current education in the home and schools actually is and what it might be, but those who have occasion to use the book as a whole will regret that Dr. Paton has not employed more uniformly the direct and luminous style of which he shows himself occasionally to be the master and that he has chosen to present so large a part of his material in so abstract a manner. A few detailed cases would have been both more interesting and more illuminating.

EDMUND C. SANFORD.

Developing Mental Power. By STRATTON, GEORGE M. Boston, Houghton Mifflin, 1922. 77p. \$.80.

Appearing as one of the Riverside Educational Monographs this book is patently a book on Educational Psychology for teachers, treating of the questions, what is mind, and how shall it be trained?

The opposing views of extreme faculty psychology and extreme particularization of function are presented, though naturally the supporting arguments of each could not be very largely developed within the compass of seventy-seven small pages. The educational implications of each view are treated by means of reviews of the discussions centering around transfer of training. The conclusion is reached that certain habits can be generated through school discipline, habits of giving "controlled attention to the task in hand; energetic attack upon it; accuracy in interpreting, remembering and reporting what is seen or read or heard; the power to distinguish important and unimportant. These are part of intellectual training; these and other things take the place of the few faculties of the older belief." Some of the other things are, emotional and volitional control which the author gives an important place in the educational process, even going so far as to suggest exercises in "will training" which the teacher can give the child: "Parents and teachers might well invent and assign things to be done, rewarding in themselves, and chosen, perhaps, from cooking,

drawing, modeling, painting, acting, reading, or any other of a hundred things—but now used in order to make habitual the right ways of purpose, applicable in any work. These right ways might here be set down, with another purpose than was guiding us earlier in this section, as: (a) suitable forethought; (b) speed and energy of attack, once the decision is made; (c) perseverance in what is undertaken; (d) economy of action, elimination of waste effort, “form”; (e) excellence of result in the product; (f) restoration of order when the work is done, putting away of tools and materials, clearing and cleaning up. Each of these six phases of the process should receive due attention,—perhaps one at a time, as Benjamin Franklin practiced the virtues,—but recurring and with different degrees of difficulty. There should be brief explanation before and after the fact, that the *idea* of what is sought should come with the practice, and should help to make the practice itself more fruitful and ready to reappear spontaneously in new places. And whatever is approved elsewhere as a means to interest and progress might be used here; if “marks,” rewards, praise, or rivalries are good to spur on in numbering or writing or any other study this present learning to will aright is as worthy of their incentive.

As an exercise in suitable fore thought, the following might serve as a door to something better. There is, let us say, but ten minutes left, and the child must choose between cutting some design in paper and making candy; and the choice is then appraised, with explanation, according as the child has stopped to think, to look ahead, before deciding. Or, again, having at hand only some modeling wax, a pair of scissors and some very narrow strips of thin colored paper, one must decide whether to build a paper house or make the figure of a dog. Or still again, the child, without actual materials at hand and with the use only of his imagination, must say—with no change of vote permitted—which line of conduct is suitable, either in cases like those just given, or where some one of a thousand other situations is described—where, perhaps, a child has visiting playmates who have come walking from afar and up a long steep hill; shall they at first play “authors” or play “tag”?

Other evidences that the author has not been able entirely to avoid the extravagances incidental to popular psychologizing are such sentences as: “Steadiness [of will] not only has ten times the effect of violence, it is ten times more readily attained.”

Nevertheless the monograph presents some considerations that may be new and instructive to young teachers even in these days when nearly everyone who goes to normal school takes a course in elementary or educational psychology, and by its stress upon the need of experimental verification of educational practices it should perform a service, to teachers, to psychology and to education. H. D. KITSON.

Psychotechnik und Taylor—System. K. A. TRAMM. Berlin, 1921, mit 89 abbildungen, VII, # 139.

Critics of the Taylor System have pointed out time and again that it has overemphasized the mechanical aspects of work at the expense of the human aspects. We know that it was certainly not Taylor's intention to do this, but that he was hampered in his approach to the problems of human behavior by the lack of an adequate technique.

Most of his followers were even less well equipped in this respect than he, and consequently, the subsequent development of Taylor's system was increasingly lop-sided. We have, as a result, the technique

of time-study and task setting which not only gives inadequate consideration to the human aspect of work but which is full of fallacies from a statistical point of view. Notable exceptions to this tendency are the Gilbreths and Robert Wolf. The latter called attention to the beneficial effects of giving to individuals graphic information about the progress of their work. The former attacked the more fundamental and comprehensive problem of developing the best possible methods of doing work by studying tasks in terms of the activities of the worker by which they must be carried out. Various methods were used to accomplish this, the best known being the photographic motion study.

In this country, the fundamental soundness of the ideas underlying the Gilbreth technique has been ignored because of its mechanical intricacies, so that the Gilbreths are today somewhat outside the pale of the more orthodox followers of Taylor. On the other hand, on the continent, and particularly in Germany, where the Taylor system has at least as great a vogue as in this country, Frank Gilbreth is regarded as its foremost exponent. And consequently, the development of the Taylor system there has been, at almost every point, influenced by Gilbreth's emphasis upon the study of human behavior and human capacities. The extensive literature on scientific management which the Germans have developed is an eloquent testimony of this influence, for it is difficult for any one to pick up anything on the subject of Taylor without finding reference to Gilbreth or accounts of his work either there or in this country. And the extensive work which the Germans themselves have done has definitely taken the direction of a study of the problems surrounding the performance of industrial tasks.

This is particularly true of the German psychologist. Whereas in this country, psychologists have given little or no consideration to Gilbreth's work, even though his technique is the ideal embodiment of the behavioristic point of view, in Germany it has been the stepping stone by which they have transplanted their methods directly into industry. And consequently, while applied psychologists in this country have been dabbling with the fringes of industry, German psychologists have stepped directly into its heart.

With this introduction, the significance of Tramm's title "Psychotechnik und Taylor System," as well as the contents of the book, can be more readily understood. Tramm is one of a group of industrial engineers and psychologists among whom are W. Moede, G. Schlessinger, A. Wallich, S. Sinner, Alex. Schackwitz, E. Mickel, C. Piorkowski, who have been active in this field. Tramm begins with an outline of the general factors which apply to all work, such as physical environment, work periods, living conditions and habits of life, individual and team work, supervision, and the elements into which all these factors are subdivided. He discusses them merely in order to indicate their importance, without any attempt to show how they are to be controlled. Most of his book is devoted to the discussion of experiments and methods calculated to show how particular occupations can best be arranged with reference to the worker. For example, a study of about one hundred tram motor-car men under experimentally controlled conditions resulted in a modification of the handles of the three levers which they were required to use and also in a standard method by which these levers could be manipulated to the best advantage. Throughout the book, emphasis is placed not only on the details or steps by which an occupation can be best performed,

but upon the construction of the tools and apparatus peculiar to it as well. The description of the laboratory devices for measuring the performance of individuals brings out vividly the contrast between the German approach and that of the followers of Taylor in this country. Instead of the lonely stop watch from which time is recorded only by a fallible observer, Tramm describes in addition a half dozen other devices of varying degrees of refinement, devices which record operations graphically and which can be adapted to a wide range of occupations. Gilbreth's motion study technique is also fully discussed, though without any explicit reference to its originator.

The effects of practice are investigated with special reference to the process of training workers. And one of the outstanding features of the book is the presentation of a pictorial technique for instruction purposes. The study of fatigue is probably the most unsatisfactory section of the book, a fact which can be readily understood when we know the nebulous stage in which the fatigue problem rests today.

The psychologist will undoubtedly find much to criticize in this book. Its experiments are not presented with sufficient facts or data to enable him to estimate the value of their results. There is a too ready acceptance of some of the traditional and worn out concepts of psychology. Inferences are drawn which are often unwarranted by the facts. And the academic or superficial point of view is still very noticeable. Industrial engineers on the other hand may well question the practicability of the technique and the value of some of the results. Nevertheless, the book as a whole is an admirable attempt to visualize the problems of production, not only in relation to output but in relation to human capacities and abilities as well. As such it is an excellent critique of Taylorism as it exists in this country and should be of considerable suggestive value to psychologists, industrial engineers, and others who are interested in industrial practices. HENRY C. LINK.

NEW BOOKS RECEIVED

Books for review should be sent to W. F. Book Department of Psychology, Indiana University, Bloomington, Ind.

FREDERIC PIERCE. *Our Unconscious Mind and How to Use It*. E. P. Dutton and Co., New York, 1922. 323 pp. Price, \$3.00.

ARTHUR S. OTIS. *Otis Self-Administering Tests of Mental Ability: Higher Examination*. World Book Company, New York, 1922. (Form A and Form B; price per package of 25 examination booklets, \$1.00.)

Though the manual for administering this test is still in mimeograph form the tests are ready for use and the mimeographed sheet of directions contains all instructions for the successful use of the examination.

The publishers announce that the Higher Examination will be followed by an Intermediate Examination for grades four to eight. There will then be available self-administering tests for use in testing any literate person.

J. L. STENQUIST. *Stenquist Mechanical Aptitude Tests*. World Book Co., New York, 1922. Test 1, Pictures of common mechanical objects. Test 2, In part similar to test 1 and with questions relating to machines. Price per package of 25 examination booklets (either test) including one key and one record sheet, \$1.50.

- Manual of directions, 21 pp., contains instructions for giving, scoring and interpreting results and norms for each test. Price, \$20.
- — —. *The Reorganization of Mathematics in Secondary Education*. U. S. Bureau of Education, Bulletin, 1921, No. 32, Washington, D. C. 73 pp.
- — —. *Summary of Conclusions and Recommendations from a Survey of the Public Schools of Elizabeth City, N. C.* Department of the Interior, Bulletin No. 26, 1921, Washington, D. C. 43 pp.
- — —. *Salaries of Administrative Officers and Their Assistants in School Systems of Cities of 25,000 Inhabitants or More*. Department of the Interior, Bulletin No. 30, 1921, Washington, D. C. 38 pp.
- HENRY S. CURTIS. *School Grounds and Play*. Department of the Interior, Bulletin No. 45, 1921, Washington, D. C. 31 pp.
- EDGAR DAWSON. *Preparation of Teachers of the Social Studies for the Secondary Schools*. Department of the Interior, Bulletin No. 3, 1922, Washington, D. C. 24 pp.
- HERMILIO VALDIZAN Y HONORIO F. DELGADO. *Revista de Psiquiatria y Disciplinas Conexas*. Publicacion Trimestral. Lima, Peru, C. F. Southwell, 1921, Vol. IV, No. 1. 72 pp.
- — —. *Record of Current Educational Publications*. Department of the Interior, Bulletin No. 52, 1921, Washington, D. C. 32 pp.
- — —. *Tijdschrift Van Den Nederlandschen Werkloosheids-Raad*. Bureau: Singel 530, Amsterdam. Vol. 5, No. 2, 1922.
- MYERS, CAROLINE E. and GARRY, C. *Measuring Minds*. An Examiner's Manual to Accompany the Myers Mental Measure. Newson and Co., New York and Chicago, 1921.
- MCCALL, WILLIAM A. *How to Measure in Education*. The Macmillan Company, New York, 1922.

THE JOURNAL OF APPLIED PSYCHOLOGY

VOL. VI

SEPTEMBER, 1922

No. 3

HORMONES AND INTELLIGENCE

(CORRELATION BETWEEN THE RATIO OF HEIGHT TO
WEIGHT AND INTELLIGENCE, IN NORMAL
INDIVIDUALS.)

By Sante Naccarati, M. D., Sc. D., Ph. D., and R. L. Lewy-Guinzburg, B. A. (From the Department of Psychology,
Columbia University.)

The physiological experiments concerned with the endocrine glands and the sympathetic nervous system, during the last few years have attracted the attention of the psychologist to a field which at first seemed to be only a branch of Medicine.

At that time when James announced his theory on the emotions practically nothing was known about the relationship existing between the internal secretion glands, the sympathetic and the central nervous system. But today, anatomy, histology, and physiology have made great progress in the study of this subject, bringing to light facts, which are not without interest in the realm of psychology. Such conceptions as the sympathetic nervous system being the seat of the subconscious life, the hormones of some internal secretory glands being responsible for the formation of characters, temperaments and for the emotions, the thyroid being the gland of the intelligence, and recently the relationship found between the derangement of the endocrine-sympathetic system and the war neuroses, are facts, which, even if not yet accepted by all, open a great field for psychological experimentation.

The hypothesis recently advanced by an Italian investigator (*Pende*) that the same hormones which promote the morphogenesis of the skeleton, and muscles of the limbs, promote also the development of the psychomotor and psychosensory centers, would lead to the conclusion that there exists a correlation between the morphological development and the intelligence of the individual.

Hormones is the term given by *Starling* (who took it from the Greek *ormao*, to stimulate, to excite) to the products of

internal secretion of the endocrine glands, namely; thyroid, suprarenals, parathyroids, thymus, gonads, hypophysis, epiphysis, etc., which are able to act on organs distant from their origin by provoking physiological actions.

These physiological actions have been grouped in three categories, namely:

- (a) Those which regulate the morphogenesis of the body or morpho-regulators.
- (b) Those which regulate the metabolism or chemio-regulators.
- (c) Those which regulate the nervous reactions or neuro-regulators.

If the forms of the body, the metabolism and the nervous reactions are under the control of the endocrine system, then we can assume that the individual constitutions and temperaments are nothing else than the particular anatomical conformation and the individual manner of functioning of this system.

The constitutions as dynamic factors in individual differences are usually not considered with other factors such as age, sex, environment, by students of human behavior. Nevertheless they play an enormous rôle.

Quite recently one of the writers (1) has shown the rôle that the morphologic element plays among the predisposing causes of the psychoneuroses. It is within the endocrine glands that reside the *endogenous* causes of the individual differences, namely the causes which are located in our organic machine and which are very active during the period of development of the organism.

It is known that among the different hormones exist functional correlations which are still the object of study and discussion. Upon such correlations of mutual help, of reciprocal inhibition or functional anatagonism depends what is called "*interhormonic equilibrium*." Variations of the interhormonic equilibrium in favor of one or another category of hormones account for the psysiological and pathological deviations of development and behaviour.

We do not as yet possess a complete knowledge of the hormones of the different glands, nor do we know in a positive manner all the actions of those hormones whose existence have been demonstrated. The hormone which has been better studied is adrenin, which is secreted by the medullary portion of the suprarenal glands, and whose secretion according to experiments of Cannon, (2) is greatly influenced by some emotions.

Falta (3), Biedl (4), Pende (5), have divided all the hormones into two classes; in one class are included all the hormones which promote and in the other all those which oppose the triple actions

which we have spoken of. They use respectively the terms *accelerator*, *disassimilator*, *catabolic* hormones and *retarding*, *assimilator*, *anabolic* hormones. Among the hormones of the first class or category are included adrenin, some hormones of the thyroid, some of the hypophysis and some of the gonads. *Pende* has made a further distinction of hormones promoting the development of the animal system (constituted by organs and apparatus which mediate contact with the external world), and hormones which promote the growth of the vegetative system.

Referring to the works of *Viola* he points out that the development of the visceral system and that of the animal system show a certain independence, even antagonism, in the sense that the two great systems (differentiated by *Bichat*) do not grow simultaneously, but in alternate phases: and the more an organism develops the vegetative system, the less it develops the animal.

Starting from the principle, *Viola* (6) differentiated three morphologic types of individuals:

(a) A *microsplanchic* type, in which the excessive development of the limbs predominates over the relatively deficient development of the trunk.

(b) A *megalosplanchic* or *macrosplanchic* type in which the exaggerated development of the trunk predominates over the relatively deficient development of the limbs.

(c) A *normosplanchic* type in which there exists a harmonious proportion in the development of the trunk and the limbs, because neither one predominates over the other.

So the morphologic type is determined by the relation or proportion between the development of the trunk, which includes the system of nutrition or vegetative system, and the development of the limbs, which include the organs which mediate contact with the external world (animal system).

The development of either one of these two great systems, is, according to the hypothesis of *Pende*, under the protective influence of a special category of hormones. "The harmony of the individual forms is dependent upon the steady and rhythmic function of these two categories of hormones." He thinks that the group of hormones which is called *accelerator*, *catabolic* or *disassimilator*, favours the morphogenesis of the skin and its annexes, of the skeleton and the muscles of the limbs, and the development of the psychomotor and psychosensory centers.

Of course this hypothesis was based on clinical and speculative grounds. Its demonstration would be of far reaching importance to the psychologist and educator, inasmuch as it would mean that

individuals in whom the group of accelerator or catabolic hormones predominates over the other group of retarding or anabolic hormones, are intelligent; and that the development of the intelligence may be favored by stimulation of the catabolic group of hormones. This group is constituted mainly by the hormones of the thyroid, and also by adrenin and by some hormones of the sex glands and of the post-pituitary. As shown in a previous work the individual possessing catabolic hormones in excess correspond to the hyperthyroid and microsplanchnic type; therefore a positive correlation should be expected between this morphologic type and intelligence, in the sense that the intelligent people should be found preferably among individuals showing a dolicomorphic tendency, if the hypothesis of *Pende* were true.

One of the writers (7) starting from the morphologic researches of Viola, who demonstrated that the microsplanchnic type is a hyperevolute type, gave experimental support to these views by showing that a positive correlation exists between intelligence and either the morphologic index or the ratio of height to weight.

The morphologic index of a given individual is obtained by dividing the linear value of the sum of one upper and one lower extremity by the volumetric value of the trunk. The method for the estimation of the morphologic index is given in full in the publication just quoted.

In this supplementary research we are given the results of the correlation found between intelligence and ratio of height to weight in three groups of university students, two of which were already mentioned in the same work.

The first group comprises 50 students mostly female who were given the Army "alpha" test. The correlation of this group which gave $+ .44$ and a P. E. of 0.06 is reported in table 1.

The second group is made up of 94 university students of both sexes to whom the "Otis" test was administered. The correlation of this group which gave $+ .15$ and a P. E. of 0.07 will be found in table 2.

The third group comprises 252 male college students all below the age of 21 to whom the Thorndike Examinations test was administered. The correlation of this group was $+ .13$ (Pearson formula) with a P. E. of 0.041 . We are indebted to Prof. Thorndike and Dr. Wood for this group.

Probably the smaller coefficient found in this group is due to the fact that it included too many normosplanchnics. In order to obtain larger coefficients of correlation a group should be made of subjects who offer large morphologic and intelligence differences,

viz: mainly microsplanchnics and macrosplanchnics and individuals who did not receive college or high school education.

Individuals above the age of 25 are not suitable subjects for this kind of investigation, as it is known that after the age of 25 many external factors, aside from the internal ones that are under the influence of the ductless glands, may influence the weight of the body, such as occupation, kind of life, marriage, etc. We were compelled to overlook these factors, in the case of the first two groups, because of the difficulty of obtaining at the time a large group of subjects of the same age and sex, although we were aware of their importance as causes of error.

The association of male and female subjects in the same group must also be responsible for errors and deviations in table 1 and 2. For these reasons and because of the small number of subjects the results obtained from groups 1 and 2 must not be taken at their face value.

In table 2 the correlation between height and intelligence and between weight and intelligence gave smaller coefficient than the correlation between ratio of height to weight and intelligence (respectively $+.02$ and $+.11$).

Still lower were the coefficients found in table 3, which owing to the larger number of cases, to the standardization of sex and age and to the better intelligence tests used, seems to us quite significant (r Intelligence Height $=-.11$; r Intelligence Weight $=-.02$) if we consider also the results obtained in the previous work already quoted (r 221 subjects Intelligence-Ratio of height to weight $=.23$, P. E. $=.04$; r Intelligence-Height $+.04$; r Intelligence-Weight $-.18$).

In table 1 we found the coefficient of correlation between weight and intelligence is higher than the coefficient between ratio of height to weight and intelligence, although the coefficient of correlation between height and intelligence was lower ($+.14$).

Technical difficulties compelled us to take the measurements in inches and pounds instead of in the metric system, as it was done in previous work. We did not think necessary to transform the measures obtained in inches and pounds in their equivalent in metric system as our interest lay in the correlation not in the individual values. A fairly approximate transformation of the metric indices given in inches and pounds, and vice-versa may be obtained, or multiplying, such indices by 5.58.

The ratios of height to weight found in the previous group of 221 college students, range from about .370 to .639 (as calculated

from the equivalent values in metric units in which they were measured, viz: from 2,066 to 3,565).

In the subjects given in these groups the ratios range from .321 to .640, which correspond approximately to 1,791 and 3,571 in the equivalent metric values, the larger range of ratios being due to the fact that the subjects used in these groups have greater age limits.

The correlation between the morphologic index and the ratio of height to weight as given in the previous work is $+.7$ or more.

It must be remembered that the ratio of height to weight is merely an approximate correspondent value of the morphologic index. To obtain better results in the correlation the morphologic index should be used instead of the ratio height to weight.

TABLE 1

NO.	AGE	SEX	HEIGHT INCH	WEIGHT LBS.	RATIO INCH H:W LBS.	ALPHA TEST SCORE
1	21	F	65	110	.590	176
2	21	F	62	112	.553	174
3	39	F	62	111	.559	173
4	17	F	68.5	122	.561	170
5	17½	F	63	114	.552	168
6	19	F	63	116	.543	162
7	30	F	66.5	121	.550	160
8	22½	F	60.2	128	.470	158
9	17	F	62	120	.516	150
10	17	F	65	120	.545	147
11	20	F	64	126	.508	144
12	20	F	63	108	.583	139
13	36	M	71	160	.443	138
14	18	F	67	127	.519	137
15	19	F	66	118	.558	136
16	22	F	64	139	.460	135
17	22	F	67	110	.609	133
18	18	F	64	125	.520	132
19	18	F	66	135	.488	131
20	35	F	66	175	.378	130
21	19	F	66	136	.486	129
22	21	F	63	123	.512	127
23	18	F	65	131	.496	120
24	22	F	67	140	.480	119
25	23	F	61	119	.511	118
26	19	F	63	135	.470	117
27	19	F	64	110	.581	116
28	17	F	61	103	.592	114
29	18	F	65	136	.477	111
30	18½	F	63.5	130	.488	110
31	22	F	66	139	.475	109
32	42	M	73	155	.470	108
33	20	F	62	129	.480	107
34	21	F	60	108	.555	106

CASE NO.	AGE	SEX	HEIGHT INCH.	WEIGHT LBS.	RATIO H W	OTIS TEST SCORE
35	29	F	66	175	.377	105
36	36	M	66	176	.375	(102)
37	19½	F	65.25	139	.469	100
38	19	F	64	120	.533	99
39	21	F	66	130	.507	98
40	30	F	66.5	125	.532	97
41	19	F	67	124	.540	94
42	25	F	64	126	.508	93
43	20	F	67.5	138	.489	92
44	19	F	62.25	130	.477	91
45	18	F	65	137	.474	90
46	36	F	69.5	156	.449	88
47	17	F	62	123	.504	86
48	19	F	64	125	.515	85
49	23	F	63.5	140	.453	79
50	31	M	69.5	145	.479	(76)

TABLE 2

CASE NO.	AGE	SEX	HEIGHT INCH.	WEIGHT LBS.	RATIO H W	OTIS TEST SCORE
1	30	M	71	170	.418	210
2	19	M	70	165	.424	220
3	21	F	62	132	.470	210
4	26	M	65.5	122	.537	214
5	30	F	67	127	.528	196
6	30	F	67	115	.583	203
7	34	M	73	170	.429	213
8	33	F	66	123	.537	209
9	25	F	68	138	.492	212
10	30	F	67	150	.447	189
11	22	M	69	165	.418	198
12	31	M	68	145	.469	198
13	48	F	64.5	110	.586	217
14	24	F	64	116	.552	209
15	24	F	67	127	.528	193
16	23	F	61	125	.488	193
17	36	M	70	165	.418	201
18	25	M	67	125	.536	207
19	25	F	67	115	.583	201
20	40	F	66	135	.489	190
21	31	F	64	170	.376	210
22	27	M	65.5	148	.443	193
23	37	M	69	126	.548	179
24	43	M	69	125	.552	182
25	28	F	60.5	99	.611	177
26	31	F	66	150	.440	185
27	23	F	67	168	.399	196
28	27	F	62	109	.569	186
29	20	F	63	115	.548	173
30	35	F	62	110	.564	185
31	29	F	62.5	125	.500	189
32	34	F	63	130	.485	195
33	28	F	67	163	.411	175
34	34	M	68	138	.492	148

CASE NO.	AGE	SEX	HEIGHT INCH.	WEIGHT LBS.	RATIO H W	OTIS TEST SCORE
35	26	F	67	152	.507	144
36	31	F	63	110	.573	204
37	36	F	66.5	175	.380	157
38	19	M	67.5	135	.500	201
39	40	F	63	175	.360	181
40	40	M	66.5	150	.443	186
41	19	F	65	125	.520	202
42	18	F	67.5	135	.500	194
43	20	M	69	140	.493	192
44	22	F	63.5	115	.552	166
45	19	F	67.5	139	.485	200
46	41	F	70.5	170	.415	169
47	21	M	70.5	135	.522	121
48	27	M	67.5	155	.436	202
49	21	M	70	166	.421	173
50	19	M	68	140	.486	151
51	18	F	66.5	140	.475	173
52	22	M	70	160	.437	206
53	19	M	72	160	.450	210
54	20	M	69.5	156	.445	163
55	20	M	71	162	.438	165
56	17	M	70.5	235	.300	191
57	20	M	66	130	.508	154
58	25	M	70.5	130	.542	169
59	54	M	72	175	.411	157
60	22	M	70	168	.417	171
61	21	M	69	153	.433	152
62	18	M	68	170	.400	144
63	19	M	69	140	.493	177
64	21	M	73	200	.365	187
65	20	M	69	138	.500	157
66	30	M	64	141	.454	154
67	20	M	65	167	.389	151
68	18	M	69	178	.387	184
69	30	F	64	150	.427	165
70	22	F	64	120	.533	162
71	30	F	67	120	.558	177
72	32	F	63.5	132	.481	173
73	26	F	67.5	160	.422	142
74	19	M	69	154	.448	205
75	26	M	65	130	.500	185
76	19	M	65	145	.448	133
77	28	M	67	150	.447	207
78	21	M	67	144	.465	130
79	29	F	66.5	137	.485	205
80	31	M	66	161	.409	215
81	25	M	70	135	.518	211
82	25	M	70.5	133	.530	175
83	29	M	70	131	.543	209
84	23	M	74	145	.510	134
85	25	M	66	160	.412	203
86	19	M	61	113	.540	211
87	27	M	70	151	.464	196
88	28	M	66.5	145	.452	166

CASE NO.	AGE	SEX	HEIGHT INCH.	WEIGHT LBS.	RATIO H W	OTIS TEST SCORE
89	25	M	67.5	142	.475	204
90	26	M	71.5	146	.490	199
91	31	M	67	150	.447	178
92	37	M	66	126	.524	165
93	28	M	69	160	.481	153
94	20	M	70	158	.448	203

TABLE 3

NO.	THORNDIKE TEST SCORE	HEIGHT INCHES	WEIGHT LBS.	RATIO INCHES H:W LBS.
1	82	68	135	.504
2	87	70	138	.507
3	70	66	130	.508
4	81	69	155	.446
5	72	69	140	.493
6	104	69	130	.531
7	80	64	100	.640
8	87	70.5	125	.564
9	90	64	153	.418
10	73	64	112	.420
11	96	68	157	.433
12	66	70.5	148	.476
13	75	69	145	.476
14	68	66	130	.508
15	66	70	150	.467
16	66	64	122	.525
17	102	66	125	.528
18	65	66	135	.489
19	60	70	136	.515
20	60	70	190	.368
21	73	58	142	.408
22	70	70	155	.452
23	56	61	190	.321
24	82	66	143	.461
25	49	73	167	.437
26	78	69	150	.460
27	81	68	125	.544
28	74	71	154	.461
29	57	62	127	.488
30	103	70	133	.526
31	68	74	148	.500
32	71	74	165	.443
33	63	68	123	.558
34	64	66	135	.489
35	87	68	125	.544
36	100	67	135	.496
37	66	70	165	.424
38	88	67	135	.496
39	70	70	130	.536
40	95	67	140	.479
41	61	70	160	.433
42	119	70	156	.449
43	70	71	160	.444

NO.	THORNDIKE TEST SCORE	HEIGHT INCHES	WEIGHT LBS.	RATIO INCHES H:W LBS.
44	86	72	155	.465
45	86	62	127	.488
46	72	68	158	.480
47	74	65	140	.464
48	80	69	150	.460
49	81	68	148	.475
50	92	66	138	.478
51	54	65	150	.488
52	60	70	150	.466
53	59	69	132	.528
54	72	69	130	.531
55	97	67	160	.419
56	63	68	153	.444
57	83	70	135	.518
58	85	68	150	.458
59	85	70	140	.500
60	64	66	125	.528
61	76	69	144	.479
62	90	69	140	.498
63	80	70	150	.466
64	75	69.5	145	.493
65	81	71	157	.452
66	51	72	179	.402
67	96	61	97	.629
68	60	67	125	.536
69	77	64	115	.557
70	82	72	165	.486
71	76	67	148	.452
72	93	68	145	.469
73	77	67	175	.383
74	88	66	140	.471
75	60	73	175	.417
76	81	70	148	.478
77	47	65	145	.448
78	90	63	170	.371
79	67	63	130	.485
80	58	69	155	.445
81	87	65	150	.483
82	100	68	135	.504
83	64	63	120	.525
84	51	71	168	.423
85	56	67	147	.450
86	76	70	135	.519
87	68	68	148	.459
88	65	67	138	.486
89	96	61	128	.477
90	73	68	148	.459
91	82	69	139	.496
92	77	64	131	.489
93	85	66	140	.471
94	79	65	120	.542
95	110	71	130	.394
96	60	69	135	.511

NO.	THORNDIKE TEST SCORE	HEIGHT INCHES	WEIGHT LBS.	RATIO INCHES H:W LBS.
97	76	68	148	.459
98	84	68	160	.425
99	85	71	150	.473
100	86	69	140	.493
101	55	68	142	.479
102	65	66	130	.508
103	99	71	132	.538
104	62	62	118	.525
105	80	67	135	.496
106	82	70	162	.432
107	73	68	150	.453
108	103	67.5	128	.527
109	94	66.5	126	.528
110	84	67	145	.463
111	98	65	130	.500
112	67	69	140	.493
113	67	67.5	130	.519
114	103	72	156	.462
115	60	70	155	.452
116	72	63.5	125	.508
117	76	67	125	.536
118	74	68	141	.482
119	80	68	134	.507
120	64	63	118	.534
121	72	68	135	.504
122	106	66.5	135	.493
123	65	72	150	.480
124	100	66	130	.508
125	77	63	117	.538
126	76	70	150	.466
127	69	66	130	.508
128	88	71.5	155	.461
129	62	67	140	.479
130	83	68	175	.389
131	73	69	130	.531
132	82	69	150	.460
133	80	70	150	.467
134	52	69	140	.492
135	91	66	150	.440
136	67	61	164	.372
137	72	68	148	.459
138	84	65	117	.556
139	89	67.5	138	.490
140	104	71.5	135	.386
141	71	65	115	.565
142	72	66	117	.564
143	89	68	155	.439
144	77	60	157	.332
145	47	74	178	.416
146	66	70	129	.543
147	89	64	523	.500
148	79	63	130	.523
149	95	67	130	.515

NO.	THORNDIKE TEST SCORE	HEIGHT INCHES	WEIGHT LBS.	RATIO INCHES H:W LBS.
150	73	67	123	.545
151	78	65	123.5	.526
152	63	71.5	163	.489
153	98	69	135	.511
154	85	65	140	.484
155	84	71	160	.444
156	106	65	116	.560
157	81	69.5	165	.421
158	90	66	125	.516
159	89	66	127	.520
160	60	65	130	.500
161	54	68	150	.453
162	64	67.5	175	.386
163	59	64	118	.542
164	58	68.5	138	.496
165	67	68	150	.453
166	88	68	138	.493
167	65	64	132	.485
168	92	66	150	.440
169	59	69	190	.363
170	100	68	120	.567
171	65	62	107	.579
172	68	71.5	150	.477
173	80	66	136	.485
174	78	67	140	.479
175	85	69	149	.464
176	43	71	137	.513
177	66	67	140	.479
178	76	68	130	.523
179	74	75	170	.441
180	76	69	136	.507
181	84	68	145	.469
182	72	65	130	.500
183	46	66	141	.463
184	75	67	130	.515
185	81	62	115	.539
186	77	63	125	.504
187	68	69	142	.486
188	53	69	175	.394
189	86	69	140	.493
190	69	64	124	.516
191	76	64	119	.537
192	86	65	117	.556
193	69	72	165	.436
194	95	65	115	.565
195	73	67	150	.447
196	72	66	150	.440
197	100	65	140	.443
198	64	66	132	.500
199	71	64	130	.492
200	80	63.5	110	.577
202	65	68	130	.523
203	62	69	140	.492

NO.	THORNDIKE TEST SCORE	HEIGHT INCHES	WEIGHT LBS.	RATIO INCHES H:W LBS.
201	58	68	165	.412
204	62	68	140	.486
205'	87	68	149	.456
206	87	68.5	144	.475
207	65	68	146	.466
208	84	65	157	.414
209	87	74	165	.441
210	71	68	187	.496
211	98	68	188	.493
212	57	63	103	.602
213	59	78	141	.518
214	79	64	145	.441
215	60	72	197	.365
216	65	71	145	.489
217	110	69	185	.511
218	79	70	175	.400
219	77	71.5	160	.447
220	76	72	155	.465
221	83	70	185	.518
222	81	72	159	.453
223	109	71	160	.444
224	84	67	124	.540
225	81	72	160	.450
226	66	70	187	.511
227	81	68.5	143	.479
228	72	66.5	146	.455
229	81	73	145	.503
230	104	69	188	.500
231	87	70	165	.424
232	82	65	140	.464
233	72	69	187	.504
234	94	68.5	110	.577
235	49	62	145	.428
236	64	72	169	.426
237	88	66	140	.471
238	90	69.5	162	.429
239	74	67	188	.485
240	62	67	125	.586
241	79	70	150	.467
242	56	70	160	.438
243	54	72	125	.576
244	88	68	150	.453
245	83	71	164	.433
246	76	70	145	.483
247	67	72	160	.450
248	76	71	185	.526
249	85	71	134	.529
250	96	67	125	.586
251	76	72	145	.497
252	76	64	125	.512

BIBLIOGRAPHY

- (1) Naccarati, S. The Morphologic Characteristics of Psychoneuroses.
Paper read at the Second International Congress of Eugenics, New York, September, 1921.
- (2) Cannon, W. B. Bodily changes in Pain, Hunger, Fear and Rage. New York, 1915.
- (3) Falta, W. Die Erkrankungen der Blutdrüsen. Berlin, 1913.
- (4) Biedl, A. Innere Sekretion. Berlin-Wien, 1913.
- (5) Pende, N. Endocrinologia, 2a Ediz. Milano, 1921.
- (6), Viola, G. Studi di Morfologia Clinica. Milano, 1905.
Nuovi Studi di Morfologia Clinica. Milano, 1909.
(In DeGiovanni Lavori dell' Istituto di Clinica Medica di Padova—Vol. 2, 4)
- (7) Naccarati, S. The Morphologic Aspect of Intelligence, New York, 1921.
(published in "Archives of Psychology," August, 1921.)

A COMPARISON TEST FOR INVESTIGATING THE IDEATIONAL CONTENT OF THE MORAL CONCEPTS

By R. A. Brotemarkle, University of Pennsylvania.

I. The Recognised Problem.

The pressing demand made upon Psychology today by all of the so-called Sociological Group, theorists and practitioners alike, is for an immediate statement of the Qualities of Personality and Character and an accurate set of methods for measuring the Quantity of the same in the individual. That the psychologist recognises the need is evidenced by the number of recent contributions to the field. Allport (1) has recently reviewed the contributions of the past decade, adding a selected list of seventy-one references. While reviewing all contributions he says, "Pressey's study stands out as practically the only attempt so far made to measure the complex emotional life of the individual by a simple testing device,"; which may be assumed to be his statement of the challenge of the problem. If possible might not even a simpler device be arranged?

II. The Accepted Problem.

With this demand in view the author has felt the necessity of first finding a solid psychological basis, and secondly, limiting the field to be investigated. The psychological basis accepted is the "ideational content" or direct "ideo-motor" determiner of activity. The limited field being in the general realm of "moral concepts," with such phases chosen as seemed indicative.

However great may have been the above problem or general method and field of investigation the specific method of testing bears the greatest difficulty. Whether the pitfalls of the "observation" method are to be chosen, with the resulting "variants" of the observer's judgment; whether the dangers of self-rating are to be chosen, with the resulting egotistic and compensatory judgment; whether the subject is to be informed of your intent and purpose, with the resulting "blind" or "mask"; or whether an attempt is to be made at securing a native unfolding of the personality, with the resulting rationalization, is no easy matter to decide.

With the assurance that the former methods had made no marked contributions to the problem, the latter method was accepted. The test was covered or hidden under the guise of a language test of intellectual or intelligence level with the title of "COMPARISON TEST."

III. The Test.

The form presented was printed on a page 6 3-4 x 8 inches as follows:

(1) G. W. Allport Personality and Character. *Psychol. Bull.* 1921, 18,441-455.

THE PSYCHOLOGICAL LABORATORY AND CLINIC

University of Pennsylvania

COMPARISON TEST

Name Date.....

Do not permit the common usage of words to interfere with your personal judgment. It should require but a brief space of time for working this test. Proceed quickly, but do not hurry. Your judgments must be decisive.

DIRECTIONS:

(1)

Accepting the two words already placed as expressive of the extremes, place seven words of each list in the blank spaces between them in their comparative relation.

Bad
Fair
Mean
Kind
Pure
Wicked
Considerate

Good

--	--	--	--	--	--	--	--

Evil

Cleanse
Spoil
Mend
Injure
Better
Harm
Reform

Purify

--	--	--	--	--	--	--	--

Corrupt

Brave
Timid
Daring
Heroic
Cringing
Afraid
Venturesome

Courageous

--	--	--	--	--	--	--	--

Cowardly

TESTS FOR MORAL CONCEPTS

237

Frank
Shy
Bashful
Backward
Free
Indiscreet
Sociable

Modest

--	--	--	--	--	--	--	--

Bold

Candid
Deceitful
Frank
False
Honest
Misleading
Ingenious

Truthful

--	--	--	--	--	--	--	--

Lying

Dislike
Devotion
Disgust
Charity
Regard
Contempt
Aversion

Hatred

--	--	--	--	--	--	--	--

Love

Negligent
Aspiring
Careless
Humble
Purposeful
Satisfied
Contented

Ambitious

--	--	--	--	--	--	--	--

Indifferent

DIRECTIONS:

(2)

Now return to the first test with the purpose of choosing any word in the lists or otherwise which you may think of as being worse or better, more or less forceful or more or less pleasant than the word already placed as extremes—and write them beyond the words of the margin of the page.

DIRECTIONS:

(3)

Now go over every word, including the one just written and place over each the word you may think of as better suited to express the thought which it prompts in your mind. Try to do this for every word—but pass quickly to another when you cannot.

QUESTIONS:

(1)

Now answer the following question very frankly:

"Do you feel that in this test you have revealed anything concerning your own individuality and personal character?"

(2)

If not, will you indicate those words which are expressive of your own individuality and personal character—marking with a check those which we have already uncovered or used in this test—and adding all others in a list.

(3)

Now answer the following questions very frankly:

"Do you feel convinced that your statement would coincide with the general opinion of your family, or acquaintances?"

(4)

If not, answer the following question: *"What in your opinion are the factors which make the variance between your own personal estimate of yourself and that of your family, your friends, or your acquaintances?"*

IV. Analysis of the Test.

The DIRECTIONS (1) is the attempt to secure in each of the seven series a profile of the individual concepts in such a manner as to reveal the "ideational content" of the individual, or background upon which moral conduct must be expected to take place.

Series 1-GOOD———BAD presents the background of Moral Judgment, or what we shall term "BASIC MORAL PRINCIPLE." This is the stimulus level.

Series 2-PURIFY.....CORRUPT presents the level of Moral Judgment in terms of motor response or activity, or what we shall term "THE REACTION PROFILE."

Series 3-COURAGEOUS.....COWARDLY presents the level of emotional activity or response—or what we shall term "THE EMOTIONAL RESPONSE."

Series 4-MODEST.....BOLD presents the level of assumed attitude, or what we shall term "THE EXPRESSIVE-REGRESSIVE ATTITUDE."

Series 5-TRUTHFUL.....LYING presents the level of attempted adjustment, or what we shall term "THE RELATIONAL ADJUSTMENT."

Series 6-HATRED.....LOVE presents the level of the determining emotive reagent, or what we shall term "THE UNDERLYING FEELING-EMOTIVE."

Series 7-AMBITIOUS——INDIFFERENT presents the level of final environmental adjustment or orientation, or what we shall term "THE RESULTANT ORIENTATION."

DIRECTIONS (2), DIRECTIONS (3), and QUESTIONS (1), (2), (3), and (4) are but the attempt to cast about the original results check material; first from the same strict naive approach in DIRECTIONS (2) and (3), and then from the "self-rating" approach in QUESTIONS (1) and (2), and finally from the "questionnaire" approach in QUESTIONS (3) and (4).

Care must therefore be taken to see that the note on the first page is carefully read, and that the DIRECTIONS and QUESTIONS are followed closely and answered in their order.

V. Research of the Test.

Though a definite serial ordering might previously have been determined upon from dictionary definitions or usage the standard used for college students was only arrived at after the examining of some 331 adults in college grades from Freshmen to Senior; likewise after it had been determined that no sex difference could be noted. The following are the Standards temporarily determined for college adults.

Series:

- 1 *Pure Kind Considerate Fair Mean Bad Wicked*
- 2 *Cleanse Reform Better Mend Harm Injure Spoil*
- 3 *Heroic Brave Daring Venturesome Timid Afraid Cringing*
- 4 *Shy Bashful Backward Sociable Frank Free Indiscreet*
- 5 *Honest Frank Candid Ingenious Misleading Deceitful False*
- 6 *Contempt Disgust Aversion Dislike Charity Regard Devotion*
- 7 *Aspiring Purposeful Satisfied Contented Humble Careless
Negligent*

On the basis of these standards a "deviation" scoring was made by counting the number of places or spaces which the word was

misplaced. The total possible score of "deviation" would be 128.

The following table of "deviations" is the work done on a class of men and women in the first-year course in psychology:

"DEVIATION"	MEN	WOMEN	TOTAL
8	2	2	4
10	2	2	4
12	2	6	8
14	3	5	8
16	6	12	18
18	3	11	14
20	8	13	21
22	5	7	12
24	6	12	18
26	5	5	10
28	4	6	10
30	3		3
32	3	6	9
34	7	5	12
36	6	2	8
38	2	3	5
40	1	3	4
42	2	6	8
44			
46			
48		2	2
56		1	1
60		1	1
64		1	1
Total:	70	111	181
MODE	20	20	20
MEDIAN	24	22	24

Quintiles: (Men, Women, and Total)

V 8-16

IV 18-20

III 22-26

II 28-34

I 36-64

Comparative Correlations:

Rating in Comparison Tests with Quintile in other tests: with—

1. "Total affectivity" rating by Pressey X-O Test.

26% stand in same Quintile.

28% stand in Quintile above or below - plus or minus 1.

25% stand in Second Quintile above or below - plus or minus 2.

54% rank same standing.

79% rank in standing within "tertile" limits.

2-"Idiosyncrasy" rating by Pressey X-O Test.

20% stand in same Quintile.

34% stand in Quintile above or below - plus or minus 1.

24% stand in Second Quintile above or below - plus or minus 2.

54% rank same standing.

78% rank in standing within "tertile" limits.

3-"Moral Judgement" rating by Indiana University Mental Survey Scales "Cross-Out" Tests. Schedule E. Test IV *Moral Judgment*.

22% stand in same Quintile.

43% stand in Quintile above or below - plus or minus 1.

17% stand in Second Quintile above or below - plus or minus 2.

65% rank same standing.

82% rank in standing within "tertile" limits.

4-"General Intelligence" rating by Indiana University Mental Survey Scales "Cross-Out" Tests. Schedule E. A Brief Group Scale for Measuring General Intelligence.

27% stand in same Quintile.

41% stand in Quintile above or below - plus or minus 1.

19% stand in Second Quintile above or below - plus or minus 2.

68% rank same standing.

87% rank in standing within "tertile" limits.

5-"College Grades" from College Records.

36% stand in Quintile above or below - plus or minus 1.

39% stand in Quintile above or below - plus or minus 2.

21% stand in Second Quintile above or below - plus or minus 2.

75% rank same standing.

96% rank in standing within "tertile" limits.

With a rather definite determining differentiation in a tertile form, "Good", "Medium", "Poor", a background is established by group scoring, which will enable the investigator to locate a firm foundation for his further individual analysis.

Continued research in the attempt to determine "traits" of personality, temperaments, etc. bring forth no significant group standards, but are a referable check on the individual analysis rating. It is hoped that more may be presented along this line at an early date.

VI. Discussion and Conclusions.

The above results are of a necessity of the most tentative formulation, and further analysis and verification must be made.

Several conclusions however force themselves upon the investigator in this field: 1) That under the rationalised activity of

moral comparisons and judgments there is to be found an "ideational content" or background of "moral concepts" which can be measured at least in rough tertile form as a determiner of the "intelligent", intellectual", "social", "moral", environmental adjusters of the individual.

2) That under a closer examination the individual problem of faulty adjustment might be discovered; presenting a better opportunity to the social worker, psychiatrist, chiminologist, or kindred worker to correct the underlying "motivating" background of morals, personality, or character.

3) That under the determined effort of research a test might be readily evaluated for the determining of the marks or signs of any type of personality or specific traits of character.

2 "Testing materials for Brotemarkle Comparison Test may be secured from C. H. Stoelting Co., 3037-3047 Carroll Ave., Chicago, Ill."

A METHOD FOR THE STUDY OF VOCATIONAL INTERESTS

BY MAX FREYD

University of Pennsylvania.

1

One of the goals, and not the least important one, of the worker in the field of Applied Psychology, is to harmonize the individual with his vocational environment—in other words, to provide vocational guidance in so far as the abilities and interests of the person desiring guidance are concerned. In order to determine the nature and extent of a person's abilities, we sample them by the use of standardized tests. While the materials for this purpose are of necessity limited, they are constantly increasing, especially in the direction of tests of reactions to the social environment, or personality tests. A theoretically recognized but hitherto scarcely tapped source of material for vocational guidance lies in the individual's interests. Once this source is tapped we can turn to a more intelligent study of the correlative abilities.

An individual may go into an occupation not only because he is interested in and has ability for that occupation, but also because he has a number of concomitant interests which find an outlet in the type of work he has elected to enter. Successful salesmanship is due not only to sales ability or abilities, which are difficult to determine, but also to communities of interests all of which are expressed in sales tasks; such interests, for example, as in talking to and meeting people, in variety in daily work, in walking around rather than sitting at a bench all day. It has been shown that measures of such interests are valuable supplements to tests for the selection of salesmen.

A knowledge of the interests which find expression in various occupations gives us a dynamic basis for vocational guidance and selection. Our aim should be to guide an individual into the occupation which will call into play not only his greatest abilities, but also the greatest number of his interests, in order that he may put his best energies into his work and find satisfaction in it.

It does not matter much for the purpose of test procedure how we define interests. In behavioristic terms, interest in an object may be defined simply as the tendency to make positive reactions toward that object. Introspection would add the affective element or "likes". We are not concerned in this connection with interests that are mere reflexes, such as the interest in a bright light in the periphery of vision. Rather, we are concerned with those which are developed in a social environment and are the result of heredity or social habit.

2

Very little scientific study of interests has been made up to the present. The outstanding researches are those which Thorndike has

carried out on interests in school studies. The methods used were such as to make results of limited significance, yet in their field these results are practically all that are available. They point to the conclusion that courses tend to maintain their relative degree of interest throughout school and collegiate life; that a high correlation exists between a person's interest in a school subject and his ability in it as judged by himself and that a significant, but somewhat lower, correlation exists between interest in a subject and ability in it as measured by school grades. The conclusions give added weight to any arguments for the use of interests in vocational guidance.

Interests have also been attacked indirectly from the standpoint of abnormal psychology by Woodworth¹ and Pressey². The latter's contribution presents two aspects of interest: (1) the blank is made extremely concise and compact by printing lists of words with directions for expressing one's attitude toward them at the top of the page, each word thus representing what would ordinarily be a complete question; (2) a method is presented for obtaining a total score. Pressey selected for his total score ten items crossed out in greater proportion by successes in school and ten others crossed out in greater proportion by failures. By subtracting each of the latter from the former, he obtained a total score which differentiated successes from failures. The overlapping of the distributions for the two groups, using the most favorable critical score, was only 10 per cent (¹¹/₁₀₀).

Under the direction of Dr. C. S. Yoakum, the staff of the Bureau of Personnel Research has devoted to interests an increasingly important part of its vocational studies. Dr. B. V. Moore³ used an interest questionnaire as one of a series of forms for experimentation in differentiating among a group of graduate engineers those who inclined toward design engineering from those who inclined toward sales engineering. The section of his interest questionnaire which proved to be the most reliable for the differentiation of the two groups was a record of occupational preferences. A scoring method was determined empirically for this blank; that is, if a larger proportion of the sales engineers than the design engineers liked a certain occupation this reaction to the occupation was assigned a score favorable to sales. The total score indicated whether a man's inclinations were in the direction of design or sales engineering. All the items entering into the total score were weighted one, no effort being made to determine the relative significance of the differences in proportion. With the most favorable critical score, there was an overlapping of only 18 per cent in the distribution curves of the total scores of the two groups.

Dr. M. J. Ream⁴ used interest blanks very much like those in-

cluded in this report, in his study of methods of selecting successful life insurance salesmen. In his study, a refinement of statistical procedure was introduced, namely, where the proportions of successful and unsuccessful salesmen who liked or disliked an item differed considerably, the ratio of the difference in proportion to its standard error was calculated. All items in which this ratio was one or more entered into the total score, all items being equally weighted. Distributions of total scores on the interest blanks showed that by this means 87.5 per cent of the men were placed correctly.

3

Some slight refinements in methods of construction of the questionnaires and in the handling of data have been made since the completion of these two studies.

In order to obtain a record of interests that would cover a large variety of subjects and yet be compact, two questionnaires were devised in the Bureau. Directions and representative items are given below. Instead of asking the questions in the form, "Are you interested in?", we have worded the directions so as to obtain an affective attitude, which in the end amounts to the same thing. These blanks were originally very long, the plan being to select and print on separate sheets the items that had been found significant for certain occupations. The lists may be made to include any number of items which the experimenter desires to use for special studies.

In the questionnaire on Choice of Occupations we have aimed to cover as wide a variety of occupations as possible within convenient limits. The occupational choices of any individual may be characterized as artistic, literary, scientific, solitary, social, mechanical, athletic, etc., according to the manner in which he checks the blank. There are 72 occupations in the list.

In the Likes and Dislikes questionnaire the items fall into three general groups: first, a series of physical attributes of people; second, a series of mental attributes of people; and, third, a series of miscellaneous items. The total number of items in this form is 129. In addition to the detailed analysis of this questionnaire, to be described later, we have considered the possible significance of each kind of symbol encircled by the subject in response to the items falling into the first two groups. From correlations with ratings, we have found that the person who tends to encircle a great many question marks after these items (the first 76 items in the list) tends also to be rated as flexible and adaptable ($r=.48$). The person who encircles a great many L's and l's after these items tends to be rated as submissive ($r=.43$), self-conscious ($r=.43$), and unadaptable ($r=.31$).

CHOICE OF OPERATIONS

After each of the occupations listed there are three symbols. Draw a circle around one of the symbols after each occupation, as follows:

Draw a circle around L if you would like doing that kind of work.

Draw a circle around D if you would dislike doing that kind of work.

Draw a circle around ? if you have no decided feelings toward that kind of work or know nothing about it.

Disregard any salary or social differences or any possible family objections. Consider only your interest and satisfaction in doing each of the kinds of work listed. You are not asked whether you would take up the occupation permanently; you are merely asked if you would enjoy that kind of work. Assume that you have the ability necessary for each of the occupations.

Be sure to draw a circle around one symbol after each of the occupations. Do not skip any.

Actor	L ? D	Mechanical engineer	L ? D
Architect	L ? D	Member of Congress	L ? D
Artist	L ? D	Musician	L ? D
Astronomer	L ? D	Newspaper reporter	L ? D
Auctioneer	L ? D	Novelist	L ? D

LIKES AND DISLIKES

Draw a circle around one of the symbols after each of the items below, as follows:

Draw a circle around L! if you like the item very much.

Draw a circle around L if you like the item.

Draw a circle around ? if you have no decided feelings toward the item.

Draw a circle around D if you dislike the item.

Draw a circle around D! if you dislike the item very much.

You may be in doubt in some cases. Use your own judgment, always encircling the symbols in the way which expresses your most frequent attitude.

Be sure to draw a circle around one symbol after each of the items. Do not skip any.

Fat men	L! L ? D D!	Golf	L! L ? D D!
Fat women	L! L ? D D!	Hunting	L! L ? D D!
Chinless people	L! L ? D D!	Taking long walks	L! L ? D D!
Energetic people	L! L ? D D!	Smokers	L! L ? D D!
Spendthrifts	L! L ? D D!	Interviews	L! L ? D D!

4

For the purpose of explaining the use of these questionnaires, we may give some typical results obtained with the following two groups: 43 seniors in Mechanical Engineering at Case School of Applied Science, and 30 Seniors in the College of Industries at Carnegie Institute of Technology. To a certain extent the curricula of the two courses coincide, but in one case the aim is to equip the individual with more theoretical knowledge and train him for a major executive position, and in the other case the aim is to provide the individual with more immediately practical information and to train him for a minor executive position in industry. Previous work on two such groups had disclosed small differences in an intelligence test and in a mechanical information test, and more striking differences in the Downey Will-Temperament Test, but the results were not sufficient completely to differentiate the two groups. It was apparent that the differences existing were more readily discerned by other means than tests. Both of the groups filled out the two questionnaires.

Our first step was to compare the two groups on the basis of tables showing the frequency with which each symbol after each item was encircled by each of the two groups. We next selected for further study those symbols in our tables which were encircled by significantly differing proportions of the two groups. Practice will enable the research worker to make a preliminary selection of the items which will show significant differences in proportion. Some items in which significant differences were found are given below.

Item	Group	L!	No. of Times Encircled				Total
			L	?	D	D!	
Fat men	Mech. Eng. 2	14	19	6	2	43	
	Ind.	2	3	16	3	30	
Very polite people	Mech. Eng. 2	15	9	13	4	43	
	Ind.	6	21	0	0	30	
Love stories	Mech. Eng. 7	27	6	3	0	43	
	Ind.	4	10	11	3	30	
"New Republic"	Mech. Eng. 1	5	34	2	1	43	
	Ind.	3	20	7	0	30	

The differences which appear to be significant are in the proportions of L's encircled after "fat men"; in the proportions of all the symbols encircled after "very polite people"; in the proportions of L's encircled after "love stories"; and in the proportions of every symbol but the first encircled after "New Republic".

The next step was to arrive at a scoring method for the questionnaires. After the item "fat men" 14/43 or .33 of the engineering students encircle the symbol L, whereas only 3/30 or .10 of the

industries students encircle that symbol. Obviously, then, in scoring the questionnaires, encircling the symbol L after "fat men" should be scored favorably to engineering. But in cases where the difference in proportion is not so large, we need a criterion for the inclusion of an item in the total scoring method. The significance of any difference in proportions may be determined by the use of the formula for the standard error of difference in proportions':

$$e_{12}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

where

p_1 = the proportion of one group encircling a certain symbol after a certain item, in our example .33 of the engineering students encircling L after "fat men."

q_1 = the proportion of the same group encircling the other symbols after the same item, or $1.00 - p_1$, in our example .67.

n_1 = the number of individuals in the group in which we found the above proportions, namely, 43.

p_2 = the proportion of the other group encircling the symbol in question, in our example .10 of the industries students encircling L after "fat men."

q_2 = the proportion of this group encircling the other symbols after the same item, or $1.00 - p_2$, in our example .90.

n_2 = the number of individuals in the second group, in our case 30.

e_{12} = the standard error of the difference (.33—.10=.23) between the proportions p_1 and p_2 .

The size of e_{12} is an index of the probability of the difference in proportions arising through chance due to the fact that we have merely sampled the total number of persons engaged in the two types of occupation; in other words, it is an index of the universality of our findings. According to Yule, if the difference in proportion does not exceed three times the error of the difference, it may be obliterated by an error of simple sampling on taking fresh samples in the same way from the same material. But for practical purposes the probability of the difference being obliterated in this manner is very little less if the error is twice the difference in proportion. If the difference is equal to its error the probability is comparatively great that the difference will be eliminated by taking fresh samples. Accordingly we have accepted as a criterion

of the inclusion of an item in our final score that the difference in proportion shall be at least twice its standard error.

Applying the formula to our example, we find

$$e_{12}^2 = \frac{.33 \times .67}{43} + \frac{.10 \times .90}{30}$$

$$e_{12} = .0081$$

$$e_{12} = .09$$

The ratio of the difference in proportions to its standard error (d/e) is 2.56, which is well above our criterion for inclusion.

Not only can we determine the significance of differences in the proportions of the two groups encircling individual symbols, but we can extend the method to groups of symbols. There are appreciable differences between the two groups of students in the proportions encircling ?, D, and D! after "very polite people". If we group these three symbols, we find that of the engineering students .60 encircle one of them, and that of the industries students .10 encircle one of them. The ratio of the difference to its standard error is 5.4, so that in case an individual encircles either ?, D, or D! after "very polite people" his questionnaire is scored to that extent favorably to engineering.

The suggestion has been made that items entering into the total score be weighted in proportion to the ratio d/e. Thus, if the ratio were 3, the item would be weighted plus or minus one. We have rejected this method since it does not add anything to the discriminating power of the total score, and have given an equal weight (unity) to all significant items. Items were scored favorably to engineering, that is, if a significantly larger proportion of engineering students than industries students encircled a given symbol after a given item, encircling that item was counted +1 toward the total score. In significant items where the magnitudes were reversed, where more industries than engineering students encircled the significant symbol, encircling that symbol meant that 1 was subtracted from the total score.

Some symbols which passed the criterion ($d=2e$) were not included because of the few individuals encircling those symbols.

The items which entered into the total score and the ratios of

the differences to their errors (to the nearest digit) are given below.

Item	Symbol	d/e	Weight
Automobile salesman	?	2	+1
Bank teller	L	2	-1
Baseball player	L	2	-1
College professor	L	2	-1
Draftsman	?	2	+1
Explorer	D	2	-1
Lawyer	D	3	+1
Librarian	D	3	+1
Mechanical engineer	L	5	+1
Mechanical engineer	D	3	-1
Promoter	D	2	+1
Stock broker	D	3	+1
Fat men	L	3	+1
Fat women	L! or L	2	+1
Cripples	D or D!	2	+1
Brunettes	L	2	+1
Very polite people	?, D, or D!	5	+1
Talkative people	L	2	+1
Irreligious people	L! or L	3	+1
People who get rattled easily	D!	2	+1
Fashionably dressed people	D or D!	2	-1
Summer resorts	D	2	+1
Conventions	L!	3	+1
Love stories	L	3	+1
Detective stories	L!	2	+1
"Life"	?	2	-1
"New Republic"	L	6	-1
"New Republic"	?, D, or D!	7	+1

The distributions of total scores are given in Figure 1. The differentiation is almost complete, the overlapping being only 4 per cent. The two engineering students marked with crosses indicated in their occupational preferences that they disliked mechanical engineering. We cannot follow this lead too far, since the item "mechanical engineering" enters into the total score.

The questionnaires were rescored using only those items in which the ratio d/e is 3 or more. The distributions of total scores are shown in Figure 2. This total score shows an overlapping of 5.5 per cent of the cases, and is very little less significant than the preceding total score. However, those items in which the ratio is 2 operate to draw the groups farther apart and add to the reliability of the total score.

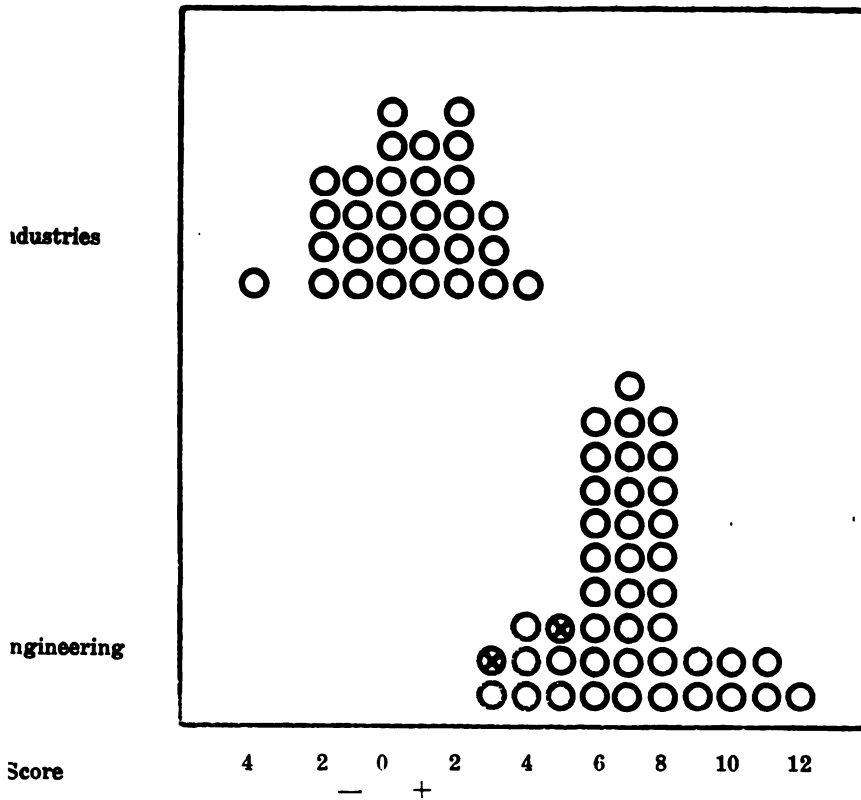


Fig. 1

Distributions of Total Scores on Likes and Dislikes and Choice of Occupations of Industries and Engineering Seniors. Items entering into the score are all weighted one. The ratio d/e is 2 or more. The two circles with crosses in them indicate engineering students who disliked engineering.

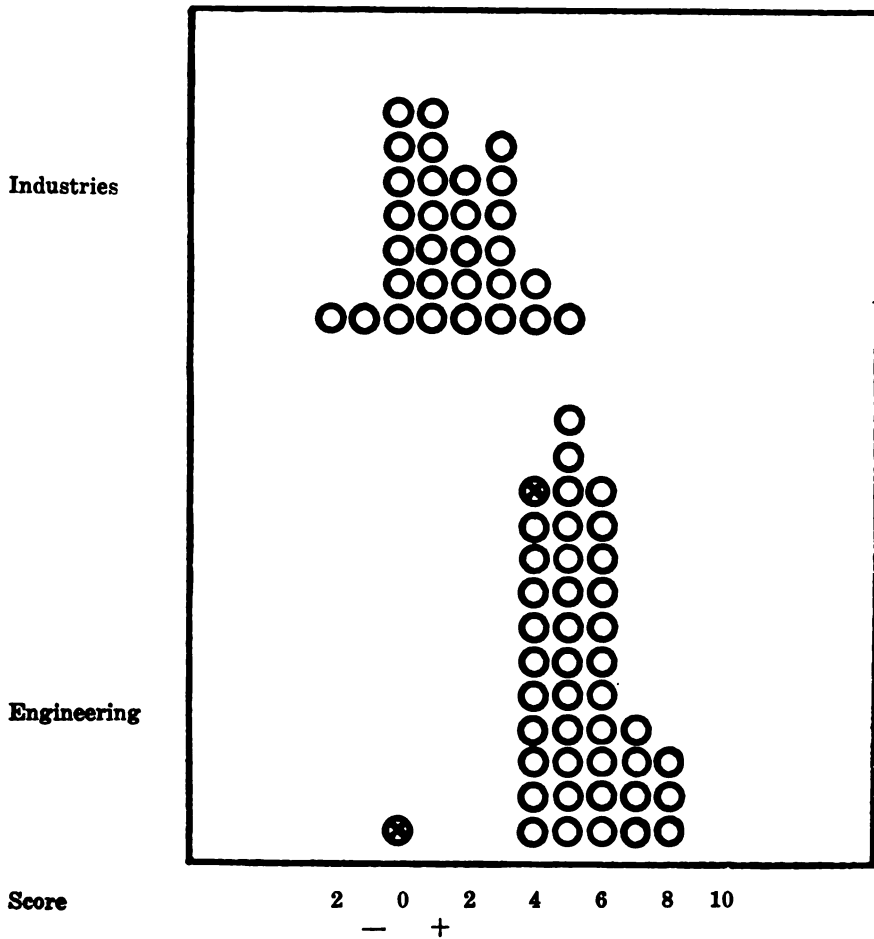


Fig. 2

Distributions of Total Scores on Likes and Dislikes and Choice of Occupations of Industries and Engineering Seniors. Items ing into the score are all weighted one. The ratio d/e is 2 or 3 or more. The two circles with crosses in them indicate engineering students who disliked engineering.

An analysis of the items that enter into the total score gives us a little insight into the character differences between the members of the two groups. With some exceptions the results would characterize the industries students, contrasted with the engineering students, as serious, quiet, thoughtful, and without particularly social tendencies except in the vocational sphere. We may perhaps conclude that the industries students are more introverted than the engineering students.

The results are not offered as absolutely conclusive evidence of the existence of such differences in the total population engaged in similar work, but point to a profitable mode of attack on these differences. No data are available on the reliability of the results obtained with this scoring method, but they will be checked up on similar groups very shortly. With the use of a similar scoring method in another connection, of eleven individuals who filled out the questionnaires twice at an interval of one month, five obtained the same score, and four others gained or lost one point over a range of twelve points. In the future use of the Likes and Dislikes blank, the L!'s and D!'s will be considered as L's and D's in devising scoring methods.

5.

The method may be briefly summarized as follows:

(1) For each of the two groups to be compared, make tables showing the frequency with which each symbol in the questionnaires is encircled.

(2) Select for statistical treatment those symbols which seem to show significant differences in the proportions of the two groups encircling them.

(3) In the cases selected, determine the differences in proportion.

(4) By the use of the formula, determine the standard error of these differences in proportion.

(5) List the cases in which the difference in proportions is at least twice as great as its standard error. These items should all enter equally into the total score.

(6) Determine which items will enter into the total score in a positive and which in a negative way. This will depend on which group will be favored in the total score. It does not matter which group we select for this purpose, but having selected one group, we should be consistent in our plus and minus signs throughout. The total score for any person will then be the algebraical sum of the positive and negative values attached to the significant items which he encircles.

In practice the procedure is not so complicated as it seems.

Several estimates and short cuts are possible and the use of the slide rule materially lessens the time of calculation.

(REFERENCES)

¹E. L. Thorndike: "The Correlation Between Interests and Abilities in College Courses." *Ps. Rev.*, 28, 1921, 374-376. Additional references are given in this article.

²See H. L. Hollingworth: "The Psychology of the Functional Neuroses", P. 118.

³S. L. Pressey: "A Group Scale for Investigating the Emotions," *J. Abn. Ps. and Soc. Ps.*, 16, 1921, 55-64.

⁴B. V. Moore: "Some Principles and Practices of Personnel Selection with Particular Reference to Graduate Engineers." *Psychol. Rev. Mon. Sup.*

⁵M. J. Ream: "The Prediction of Successful Salesmanship." Unpublished thesis, Carnegie Institute of Technology, 1921.

⁶G. U. Yule: "An Introduction to the Theory of Statistics." Formula 6, P. 269. Another formula is employed when we wish to compare the proportions of checks made by two groups, the larger of which contains all the individuals in the smaller group. We may wish, for example, to compare the interests of a group of successful salesmen with those of the total group from which they were drawn, or we may wish to compare the interests of those completing a course of study with the interests of all who started the course. The formula to be used in such cases is given by Yule: Formula 7, P. 271.

THE COLOR PREFERENCES OF CHILDREN

By S. E. KATZ
Medical School of Harvard University

AND F. S. BREED
University of Chicago

PROBLEM

The object of this investigation was to discover some of the factors upon which the color preference of children is based. The writers endeavored to determine: (1) whether certain colors are *per se* and generally pleasing, (2) whether color preference varies with increase in age and intellectual development, (3) whether the colors preferred during the adolescent period differ from those preferred during the pre-adolescent period, (4) whether social status is a determining factor in the preference for colors, (5) whether the two sexes show any difference in their aesthetic evaluation of color.

REVIEW OF PREVIOUS WORK

A similar investigation was conducted by W. H. Winch in the municipal schools of London. His results are, however, somewhat invalidated by the experimental procedure followed. He had the teacher of the children write the names of different colors on the blackboard and ask the children to write on slips of paper the names of the colors they liked best. Each pupil made a second, a third, a fourth, and a fifth choice, so that each slip of paper contained five preferences. Three spectral colors were used, and in addition black and white.

In working with school children in the present investigation, it was found that a considerable number in the lower grades were unfamiliar with even the names of the pure spectral colors. Table 1 shows the wrong names given by the kindergarten classes to the Bradley colored papers. One hundred and twenty-five children were examined individually.

TABLE 1
WRONG NAMES GIVEN TO BRADLEY COLORED PAPERS BY
KINDERGARTEN PUPILS

Color Displayed	Name Given	Number of Children
Red	Purple	1
	Brown	4
	No name	3
Orange	Red	4
	Yellow	13
	Brown	7
	Pink	1
	No name	11
Yellow	Brown	3
	Green	1
	No name	3
Green	Brown	1
	Yellow	4
	No name	1
	Blue	1
	Purple	1
	Orange	1
	Pink	1
Blue	Purple	5
	Green	1
	White	1
	Red	1
Violet	Purple	39
	Blue	38
	Green	2
	Red	2
	No name	10

The results showed that the older children were more familiar than the younger ones with the names of the elemental colors, but found difficulty in naming the different nuances of the same color. For example, objects were called red that were any color from a violet-red to a red-orange. The question may therefore be justly asked: How could Winch determine which color the children had in mind in expressing their preference for a particular color—let us say, red?

METHOD

The procedure of this experiment differed in two respects from that of Winch's experiment: Six spectral colors were used instead of three, and these colors were numbered. On a white cardboard 65 by 46 cm. in size, six rectangular Bradley colored papers, each 11 by 11 cm., were pasted in two rows, 20 cm. apart. The three colors in each row were 8 cm. apart. The colors were saturated red, orange, yellow, green, blue and violet. The number of each color was placed above it, so that in indicating their preferences there was no occasion for the children to name the colors. The test was given to all the pupils in seven public elementary schools, to all the pupils in one high school, and to several groups of college students consisting mainly of freshmen. In all, over 2,500 subjects were tested. The method of experimentation is outlined in the following directions, a copy of which was handed to the teacher together with one of the charts described above, after the latter had been fully explained.

DIRECTIONS FOR THE COLOR TEST

The teacher will please hand to each of the pupils a sheet of paper provided by the experimenter, and ask each first to write on it his name, age, grade, and the name of the school. Then the teacher will hang up the accompanying chart and instruct the children as follows:

"HERE ARE SIX DIFFERENT COLORS. EACH COLOR HAS A NUMBER. WRITE ON YOUR PAPER THE NUMBER OF THE COLOR YOU LIKE BEST."

It is very important for the success of this experiment that there be no communication between the pupils.

As soon as the test is completed, the teachers will kindly collect the papers, and return them, together with the chart and the direction sheet, to the principal's office.

RESULTS

Tables 2, 3, 4 and 5 show the total number of preferences given for each color by the pupils of the seven elementary schools, the high school, and the students of the several college classes. In Tables 2 and 3 the subjects are arranged according to school grade, in Tables 4 and 5 according to chronological age. The difference in the total number of subjects between these two sets of tables is due to the fact that some of the pupils in the grades and about one-third of the total number of college students did not give their ages on their voting cards.

TABLE 2

NUMBER OF PREFERENCES
FOR THE BOYS OF THE
PUBLIC SCHOOLS AND THE
COLLEGE MEN, ACCORDING
TO SCHOOL GRADE

Colors							Colors							
Grd. R	O	Y	G	B	V	Total	R	O	Y	G	B	V	Total	
Kdg13	12	16	13	21	4	79	9	10	17	4	13	10	63	
1	12	10	8	15	19	8	72	12	6	9	13	18	66	
2	12	8	11	14	45	17	107	13	9	11	19	60	9	121
3	18	9	5	18	50	4	99	12	5	4	13	55	5	94
4	15	10	6	11	73	8	123	14	5	7	13	82	12	133
5	21	9	7	9	62	8	116	5	3	4	15	85	6	118
6	15	3	3	13	69	4	107	5	2	3	9	75	4	98
7	13	3	2	19	64	8	109	13	5	1	18	62	6	105
8	3	0	1	4	12	1	21	4	0	0	6	15	0	25
9	9	2	4	9	17	2	43	13	1	6	6	21	2	49
10	11	0	0	5	12	6	34	1	1	4	21	8	1	36
11	4	2	1	5	12	1	25	6	3	8	11	6	0	34
12	2	0	0	2	9	1	14	1	0	2	8	13	1	25
Col. 36	15	7	58	133	36	285	49	6	20	82	64	69	290	
Total														
184	83	71	190	598	108	1234	157	56	96	238	578	133	1257	

TABLE 4

NUMBER OF PREFERENCES
FOR THE BOYS OF PUBLIC
SCHOOLS AND COLLEGE
MEN, ACCORDING TO
CHRONOLOGICAL AGE

Age	R	O	Y	G	B	V	Total	R	O	Y	G	B	V	Total
5	9	1	7	8	12	4	41	5	7	9	2	11	8	42
6	5	11	13	5	17	7	58	7	6	15	11	15	8	62
7	13	9	9	9	28	12	80	16	13	11	13	27	4	84
8	8	8	8	18	35	6	83	8	4	5	14	54	8	93
9	13	6	6	11	45	3	84	17	3	6	14	60	5	105
10	10	11	2	9	36	7	75	9	3	8	7	57	6	90
11	22	5	4	15	62	1	109	7	3	4	13	66	6	99
12	17	4	3	7	62	6	99	8	2	3	15	57	4	89
13	11	5	5	11	54	8	94	10	4	2	14	46	3	79
14	13	2	0	8	45	2	70	7	1	3	11	47	1	70
15	12	1	1	10	35	4	63	10	1	1	18	29	2	61
16	5	1	1	5	13	4	29	2	1	8	12	8	1	32
17	7	2	2	3	11	3	23	3	2	3	11	15	3	37
18	11	2	2	7	28	6	56	9	1	1	13	20	10	54
19	8	1	3	12	40	7	71	8	1	2	18	13	12	54
20	9	4	2	8	19	10	52	5	1	3	5	4	8	26
21	2	1	0	10	22	4	39	3	1	1	4	3	2	14
22	2	1	0	7	13	5	28	1	1	1	2	3	2	10
Total														
	177	75	68	163	577	99	1159	135	55	86	197	535	93	1101

TABLE 3

NUMBER OF PREFERENCES
FOR THE GIRLS OF THE
PUBLIC SCHOOLS AND THE
COLLEGE WOMEN, ACCORD-
ING TO SCHOOL GRADE

The order of color preference according to age and school grade. —In order to make a comparative study of the color preference of the different groups of subjects examined, the percentages of votes given for a color in each of the classified groups, according to school grade or chronological age, was calculated to the first decimal place and then multiplied by ten, to determine the number of preferences per thousand cases. From these comparable values the order of preference for the colors according to school grade was determined. The results are presented in Tables 6 and 7. Thus the number 1 means that the color designated by that number received the highest number of preferences, and the number 6 means that the color designated by that number received the lowest number of preferences. When two colors tied for position, the average of the combined positions is given to each of the colors.

TABLE 6
ORDER OF COLOR PREFERENCE FOR MALE SUBJECTS, ACCORDING TO SCHOOL GRADE.

Grade	R	O	Y	G	B	V
Kdg.	3.5	5	2	3.5	1	6
1	3	4	5.5	2	1	5.5
2	4	6	5	3	1	2
3	2	4	5	3	1	6
4	2	4	6	3	1	5
5	2	3.5	6	3.5	1	5
6	2	5.5	5.5	3	1	4
7	3	5	6	2	1	4
8	3	6	4.5	2	1	4.5
9	2.5	5.5	4	2.5	1	5.5
10	2	5.5	5.5	4	1	3
11	3	4	5.5	2	1	5.5
12	2.5	5.5	5.5	2.5	1	4
College	3.5	5	6	2	1	3.5

TABLE 7
ORDER OF COLOR PREFERENCE FOR FEMALE SUBJECTS, ACCORDING TO SCHOOL GRADE.

	R	O	Y	G	B	V
	5	3.5	1	6	2	3.5
	3	6	4	2	1	5
	3	5.5	4	2	1	5.5
	3	4.5	6	2	1	4.5
	2	6	5	3	1	4
	4	6	5	2	1	3
	3	6	5	2	1	4
	3	5	6	2	1	4
	3	5	5	2	1	5
	2	6	3.5	3.5	1	5
	5	5	3	1	2	5
	3.5	5	2	1	3.5	6
	4.5	6	3	2	1	4.5
	4	6	5	1	3	2

In every grade from 1 to 8, the highest number of votes was given to blue. Nearly one-half of the total number of pupils preferred that color. Green was a poor second, with red a close third. The variation in the preference for these two latter colors was much greater than it was for blue. Violet and yellow occupied the next positions, with a considerable variation from grade to grade. Orange secured last place.

When the subjects were arranged according to chronological age, blue again headed the list of preferences, but fewer groups were found voting with a majority for blue. Otherwise these tables show little variation from the previous tables. For some of the colors other than blue the order of preference was changed, but these changes were not consistent enough to justify the drawing of any important conclusions from them. As a whole, the grouping based upon chronological age showed very little difference as to color preference from one based upon school grade.

The order of color preference according to sex.—The results obtained from the female subjects are represented in Tables 3, 5, and 7. Examining Table 7, one notices again the prevalence of the preference for blue in most of the groups, but does not find this preference so pronounced as among the male subjects. In two cases it dropped to the third place, and in two to the second place. Green was the second choice, with red a close third. Above grade 9 green even moved to the first place. Violet was about equally favored by the female and male subjects. Yellow was more of a favorite with the girls of the upper grades than it was with the boys.

Table 5 shows that with the older girls blue lost in competition with green and violet. Red was somewhat less often preferred by the girls than by the boys. In general these tables do not show any striking difference in the color preferences of the two sexes.

TABLE 8
NUMBER OF PREFERENCES PER THOUSAND SUBJECTS
FOR EACH COLOR BY PRE-ADOLSCENT AND ADOLESCENT SUBJECTS

		R	O	Y	G	B	V
Pre	Male	149	83	92	133	462	79
	Female	120	79	116	122	439	151
Post	Male	156	38	27	166	501	113
	Female	134	41	72	248	394	123

Color preference in pre-adolescence and adolescence.—In order to determine whether there were any marked changes in the colors preferred after the beginning of the adolescent period, the comparable preferences for each color up to the age of fourteen for boys and up to the age of twelve for girls were added, and the sums divided by the number of groups. The same method was followed for the groups beyond these ages. The averages are represented in

Table 8. The results show that red, green, blue and violet were more frequently preferred by boys in the adolescent period than by those in the pre-adolescent period, while orange and yellow lost in value in the later period with the boys. For the girls in the adolescent period orange, yellow, blue and violet exhibited a loss in aesthetic appreciation, and only red and green a gain.

Variation in preference according to wave length.—Tables 6 and 7 show, upon closer observation, that there was a continuous rise in the values of green, blue and violet as the children advanced in age or school attainment, and a corresponding depreciation in the color values of red, orange and yellow. Table 9 shows the results obtained by adding the votes cast for these two sets of colors and dividing each total by three, the number of colors in each set. In this table red, orange, and yellow are included under colors of long wave length, and green, blue and violet under colors of short wave length. With the exception of the kindergarten classes, the colors of shorter wave length seem to have been more often preferred than the colors of the longer wave length. Only with the girls of the eleventh grade was there a tie between the two.

TABLE 9
NUMBER OF PREFERENCES PER THOUSAND SUBJECTS
FOR COLORS OF LONGER AND SHORTER WAVE LENGTH

Sex	Grade												
Kdg.	1	2	3	4	5	6	7	8	9	10	11	12	College
Long													
M.	520	416	290	324	252	320	196	165	190	531	323	280	143
F.	572	412	272	325	196	102	103	182	160	407	167	500	120
Short													
M.	480	584	710	676	748	680	804	835	810	649	677	720	857
F.	428	588	728	675	804	898	897	812	840	593	833	500	836

Variation in preference according to brightness.—By averaging the votes given to the two groups, orange, yellow and green; and red, blue and violet, the data represented in Table 10 were obtained. A glance at the table shows at once the predominance in the preference for the darker colors (red, blue and violet) over that for the lighter colors (orange, yellow, and green). Here again no sex difference was noted.

TABLE 11	TABLE 12
ORDER OF COLOR PREFERENCE FOR MALE PUPILS IN WELL-TO-DO NEIGHBORHOODS	ORDER OF COLOR PREFERENCE FOR FEMALE PUPILS IN WELL-TO-DO NEIGHBORHOODS

home training the children of the two different groups received in the very early years, the color preferences of the two groups in the early ages show distinct differences. But it is apparent from a study of the tables that the influence of this early home training is gradually overcome as the children advance in age and schooling.

TABLE 13
ORDER OF COLOR PREFERENCE FOR MALE PUPILS IN POOR NEIGHBORHOOD

TABLE 14
ORDER OF COLOR PREFERENCE FOR FEMALE PUPILS IN POOR NEIGHBORHOOD

	R	O	Y	G	B	V	R	O	Y	G	B	V
Kdg.	4	5.5	1	3	2	5.5	4	5.5	3	2	5.5	1
1	3	3	5.5	1	5.5	3	1	3	4.5	4.5	2	6
2	2.5	4	5.5	5.5	1	2.5	3	5	5	2	1	5
3	2	5	3	5	1	5	2	5.5	3.5	3.5	1	5.5
4	1.5	5.5	5.5	4	1.5	3	2.5	5	5	5	1	2.5
5	2.5	6	5	4	1	2.5	6	2.5	4.5	4.5	1	2.5
6	2	4.5	4.5	4.5	1	4.5	4	4	4	4	1	4
7	2	6	4	4	1	4	2	6	4	4	1	4

Preference for saturated colors.—As already stated, saturated colors were used in this investigation. In order to make a further study of the affective value of saturation, 250 kindergarten pupils were individually tested. Each of the six primary colors was represented by a saturated color, a tint and a shade. Table 15 represents the results obtained from this test. It shows that the boys and girls were alike in their preference for saturated colors. For the boys, the tints and shades were of equal value, while for the girls the tints were more often preferred than the shades.

TABLE 15
THE AFFECTIVE VALUE OF SATURATION ON THE PREFERENCE OF KINDERGARTEN PUPILS

	Boys	Girls
Saturated	57	70
Tint	32	42
Shade	33	25

DISCUSSION

Comparison of results with those of previous studies.—By comparing the results obtained in this investigation with those reported by *Winch*, it was found that in both cases blue received the first

position and yellow the last position in the series of color preferences. Red and green, however, received lower scores in the present investigation than those accorded to them by Winch. This might be due to the difference in procedure of the two experiments or to the difference in the nationalities of the subjects tested.

No striking difference in the color preferences of the two sexes were found, as claimed by Winch, Adams and Jastrow. That social status influenced color preference of young children, is, however, affirmed by the results presented in this paper.

In so far as the results of the present study can be compared with Jastrow's results for adult subjects and a greater number of colors, the order of preference for the colors seems to be about the same. In one other respect do the results here presented agree with Jastrow's, namely, that saturated colors are preferred to tints and shades. Washburn, Minor, and Adams reported that the less saturated colors were preferred.

It would seem that the conflicting results of the various investigations are due in no small measure to differences in the number, age, and social status of the subjects used. It is especially important to note that the number of subjects used by many of the investigators was relatively small.

Practical application of results.—Investigations have indicated that business can profit by a precise knowledge of the distribution of measures of human traits. For example, manufacturers of as well as jobbers and retail dealers in hose and shoes may be guided in production and purchase by a knowledge of the fact that measures of human feet are distributed approximately according to the curve known among statisticians as the curve of probability. In advance of lessons gained through costly experiences, the inexperienced dealer can be guided by figures based upon scientific data possessing predictive value for his trade. Naturally the same law of distribution would be of practical service to dealers in gloves, hats, shirts, collars—all articles of wearing apparel, in fact, where sizes vary according to physical measurements and the customers constitute a random selection.

In similar manner might dealers in wearing apparel for children predict in some measure the probable demand, so far as colors of materials are concerned, by such facts as those herein reported in regard to color preference. While environmental factors like social status and the season's vogue have some influence upon a child's choice of color, the fact that the blue used in this study received practically as many selections as all the other five colors together, indicates that there are certain colors that are pleasing or displeasing *per se*, a finding which has no little practical signi-

ficance. The extreme unpopularity of yellow and orange, in contrast with the popularity of blue, can be noted with equal profit.

Not only would it seem that definite quantitative knowledge of the color preferences of children could be applied in the manufacture and sale of many sorts of wearing apparel, but as well in the manufacture and sale of toys and books, and in the furnishing and decoration of children's rooms—always on the assumption, of course, that the child is permitted to select what he really wants, not required to select what his elders want him to have.

The results of this investigation suggest that efficiency experts in various lines of business could throw more definite light on the relation of color to sales by carefully planned experiments in which the actual articles varying in color would be used.

SUMMARY

1. At every age from 5 to 15 blue was most frequently preferred. Of approximately 2500 pupils tested, from kindergarten to college, 47 per cent found blue the most pleasing of the six colors used. Green was a distant second, red a close third, violet and yellow occupied the next positions and orange proved the least pleasing of the six.
2. There was in general a distinct rise in the preference values of green, blue, and violet, the colors of short wave length, and a corresponding decline in the values of red, orange, and yellow, the colors of long wave length, as the children advanced in age and grade.
3. In the comparison of pre-adolescents with adolescents, the most noticeable differences were the marked loss in the popularity of orange and yellow and the increase in the popularity of green as the children matured.
4. Red was a greater favorite among the children in poor than among the children in well-to-do neighborhoods, during their earlier years; the reverse was true of green. Differences due to social status tended gradually to be overcome as the children advanced in age and school attainment.
5. No striking differences between the color preferences of the sexes was observed.
6. Both sexes of kindergarten age showed a preference for saturated colors.
7. Definite quantitative knowledge of the color preferences of children has profitable application to the production and distribution of many articles where the child is permitted to determine the selection.

BIBLIOGRAPHY

1. Adams, H. F. Advertising and its Mental Laws, Macmillan Co., 1916, pp. 250-259.
2. Jastrow, J. The Popular Aesthetics of Color, Popular Science Monthly, v. 50, 1896, pp. 361-368.
3. Minor, A. Über die Gefälligkeit der Sättigungsstufen der Farben, Zeitschrift für Psychologie, v. 50, 1908, pp. 433-444.
4. Washburn, M. F. A Note on the Affective Values of Colors, American Journal of Psychology, v. 22, 1911, pp. 114-115.
5. Winch, W. H. Color Preferences of School Children, British Journal of Psychology, v. 3, 1909, pp. 42-65.

THE INTELLIGENCE QUOTIENT FROM TWO VIEWPOINTS

By W. T. Root.

UNIVERSITY OF PITTSBURGH

Recent literature has revealed considerable difference of opinion concerning the constancy, value and interpretation of the I. Q. Much of the discussion arises from two distinct practical situations which educators and psychologists have to meet. It is the purpose of this paper to analyze these attitudes arising therefrom and apply a little psychology to the psychologist. The writer feels that much of this difference of opinion is largely due to bias arising from the direction of attention to one's limited interests, and is more apparent than real.

The view of the psycho-statistician and the educational administrator. The psychologist hunting for statistical trends and the administrator seeking approximations find themselves in accord in the interpretation of the I. Q. They find that for the very large majority of the subjects tested the I. Q. remains constant or nearly so—sufficient for all practical purposes. For example, if a school administrator wishes to reorganize his school, let us say the Junior High, into three sections for each grade, and has the list of pupils below 80 I. Q., from 80 I. Q. to 100 I. Q., and 100 to the upper limit, he can *safely* count on these groups remaining constant for rough administrative purposes. There will be a few shiftings to be sure.

But the very weight of statistical trends and the serviceability of tests in a broad administrative way lead the practical school man who has gone in for tests to view the I. Q. as very "reliable," and he feels considerable impatience with any dwelling upon and analysis of the exceptional cases, or anything that detracts from the simplicity of the intellectual and administrative schema. Immediate practical usability in a large way dominates his thinking and action, while his daily duties accentuate this view. Again, teachers in certain sections are divided into testers and antitesters, and one may safely assume that such "all-or-none" conditions of acceptance rest on emotional belief akin to religious fervor which is likely to supplant detached scientific attitude.

Turning now to the psycho-statistician's view and the organization of the Binet tests. The Stanford Revision is so constructed that in an unbiased sampling of children the median chronological age equals the median mental age year by year. This unbiased sampling assumes of course that we have "medianaed" all factors; physical, health, hereditary, environmental, and so on. Roughly the data make a symmetrical curve, year by year, with an increasing spread of the scatter-distribution of the mental performance as we ascend the chronological age scale. It is interesting and lends

force to the constancy of I. Q. argument, that the increasing scatter fits nicely just what one would have if each variate were plotted as having a constant ratio between mental and chronological age. The data year by year or as a whole may be treated as conforming to the normal curve' (although this may not prove to be true,)—but at least it suffices for all immediate practical purposes.

Let us now assume an unbiased sampling of the child population. Will this group from year to year give a distribution similar to that obtained in the Stanford Revision (applying the usual statistical tests of central tendency and variability—median, P. E.'s, etc?) Will the mental-age rank order and the I. Q. remain relatively constant for each of the variates?

Statistical prediction depends upon the causal factors remaining constant for the mass of the variates in question. Turning now to our group under consideration: the factors of hereditary equipment which are basic will in the main remain the same, and while there will be some ups and downs, in the vast majority of cases the environmental setting will continue roughly constant. The mores, culture, ideals, aims and habits of thought will continue relatively unchanged for the different family groups. The dominant cultural change will be provided by the public schools, but considering masses, this element is injected at the same time for the same duration as in the statistical sampling (the Stanford Revision.) The burden of proof is with him who would predict that our three year group will depart seriously from the established norms as they advance year by year. Furthermore, the constancy of the factors producing intelligent behavior will maintain for the mass of the subjects considered individually, which in turn would make for progressive constancy of mental age rank-order and constancy of I. Q. It is highly improbable that any pronounced hereditary or environmental change will affect the masses in any brief period of time but it is equally true that conspicuous changes in environment, sensory deprivation or restoration, systemic conditions, emotional complexes, disease, and so on, can and do take place *in the individual case*.

The view of the clinical psychologist. So much for the ninety-and-nine that determine the character of distributions, stabilize medians, permit the psycho-statistician to indulge in scientific prophecy, and allow the educational administrator to convince his board that he has unusual clairvoyant powers. The writer confesses that his sympathies are at times with those who must meet the one, two or three that lie outside the fold of statistical con-

"Roughly symmetrical" is usually all that can be said of the data.

formity. Psychological clinics are likely to get this residue of atypical cases, consequently the clinical psychologist is less likely to be impressed by statistical trends than the educational administrator, and gets little solace from the average constancy of the I. Q. when he deals with a peculiarly exasperating subject.

This can best be illustrated with cases. Case M has had successively I. Q.'s of 93, 79, 98, and 89. The details of the case are too long to give here, suffice it to say that the same examiner administered the Stanford Revision at each examination. The subject possesses no marked peculiarities except the inability to learn to read. This inability is in the nature of almost complete failure to acquire an appreciation of phonic values. Practically all words are sight words to this subject. After the drop to an I. Q. of 79, the child was given strenuous individual phonic drill by a skilled primary reading teacher with years of experience. At the next examination the I. Q. jumped to 98. The drill was stopped for the summer vacation. On returning in the Fall the I. Q. had dropped to 89. The only explanation presenting itself at present is that the phonic drill (from which he profited at the time) improved his reading ability and hence his test performance. He has a splendid physique, teachers credit him with common-sense, good judgement, willingness and so on. The I. Q. 79 and 98 do not warrant the same prognosis, and there was nothing in the first I. Q. to lead one to anticipate the second. The environment is median and constant in character. Time and follow-up work, not I. Q., will throw light on this case. Examination of the eyes reveals no defect of sufficient importance to warrant the character of reading. If he had only accommodately maintained an I. Q. of about 60 all would have been well!

Within the last two years the Psychological Clinic of the University of Pittsburgh has handled a number of cases going into a very superior institutional home on the cottage plan, and a few cases of children going from very abject conditions into superior private homes. In the majority of these cases the institution marks a distinct improvement in intellectual, hygienic, moral and social environment. The accompanying tabulation presents a few such cases from which we have had our first report since the environmental change. It will be noted (comparing columns 4 and 10) that in all the cases except number 2 there is a distinct increase in the rate of gain in mental age following the change in living conditions. The remaining cases hug the lower level of median performance, had poor environmental conditions and are accruing mental months at a much faster rate than before environmental change. There is nothing in the cases just cited which does vio-

lence to any of the known facts concerning the origin or composition of the data from which the I. Q. is derived. Such data depend, in the by and large, upon all of the factors lying back of complex human behavior remaining relatively constant; in the above cases we have introduced marked change, tearing the child away from family habits, mores, and influences *that under ordinary laissez faire conditions might have required three or four generations to effect*. One may thoroughly believe in the constancy of the I. Q. and with perfect consistency refuse to accept statistical trends if there is evidence of marked departure between the first and second test in any one of the factors making up the composite nature of behavior. Any other procedure would be to have the statistical tail wag the causal dog.

Cases transferred from inferior homes to superior institutional life

1	2	3	4	5	6	7	8	9	10
Case	Chronological Age at time of 1st examination	Mental Age 1st examination.	I. Q. 1st examination	Chronological Age at time of *2nd examination	Mental Age 2nd examination	I. Q. 2nd examination	Time—in months between 1st and 2nd examination	Number of months gained in Mental age between 1st and 2nd examination	Column 9 ÷ 8 giving I. Q. in terms of rate of mental increase change of environment.
1	4-0	4-8	116	5-7	6-10	123	19	26	136
2	9-9	11-4	116	11-4	12-1	107	19	9	47
3	6-5	5-6	84	7-7	7-0	90	16	18	112
4	7-7	7-0	92	8-10	9-4	106	14	23	200
5	6-2	5-4	86	7-3	7-0	97	14	20	143
6	14-0	12-5	89	15-0	13-11	98	12	18	150

*2nd examination after change of environment in every case.

One case more, which the writer had up for consideration recently. The subject, a little girl, was eleven years five months old at the time of testing, I. Q. 67. (Has had simply unbelievable social conditions. Mother and father died of "flu," went with brother and sister, to live with grandfather and grandmother. Grandmother died. Children then alone with grandfather who is eccentric, morose, grossly ignorant and surly. No companions, and school several miles away....No shoes, no underwear during winter months. Only a few months schooling.No toys, no outside contacts. Squalor and poor food.) This child was as timid

as a deer. Although she was almost without vocabulary, and could not read, her ability in the Porteus maze tests, Healy-Fernald construction boards, and her complete lack of the "so-called" stigmata of feeble-mindedness pointed towards a diagnosis and prognosis more favorable than the I. Q. would suggest. Now the institution having the child examined did not want her if feeble-minded. The matter was finally presented to the institution frankly stating the dilemma, namely the highly atypical character of the environment. The institution accepted her. At the present moment of writing, she has within a few months acquired sufficient skill to read the passage from the Stanford Revision beginning, "New York, Sept. 5 . . .," without error, but still too slowly to score plus.

What prognosis should be made? To begin with the case is not quite analogous to the ones cited in the accompanying tabulation, for in these the I. Q. in no case takes the child out of the normal group. But may it not be true that the monstrous conditions under which the child has lived has markedly deflated her I. Q.? Is it not possible that if the institution in which she is placed offers her even median social milieu that the I. Q. may rise appreciably? The writer would venture the prediction that the subject may attain an adult intelligence of low normal (I. Q. of 85 or 90.)

Certain practical applications and suggestions can be made from these two contending viewpoints, statistical trends on the one hand and analysis of the qualitative factors peculiar to an individual case on the other:—

1. Probably no one not even the most enthusiastic eugenist would deny that in trends and for school administrative purposes in the large, the results of an examination of a school either by the Stanford Revision or a battery of mass tests must be accepted as highly indicative of relative rank-order of the intellectual status of the group studied, and *vastly superior to any impressionistic estimate.*

2. More than this, in the individual case of fluctuation of I. Q. should arouse the attention as to the cause. The first assumption should be that a clerical error in scoring has been made, or age has been reported differently at different times. This "hunch" is often correct. (We are assuming trained examiners throughout.) Up to this point the burden of proof is with the case that departs markedly from statistical trends. A search for causal factors and an analysis of the nature of the response must now be made. But having verified the accuracy of the data no apology needs be offered to statistical methods; neither is it

in order for statistical method to offer an apology for failure in individual conformity.

Statistical method was developed to anticipate trends in data and error therein. At this point in our thinking a quotation from a recent article by Yule¹ is well worth pondering:—

“Statistical methods, I say, should be regarded as ancillary, not essential. They are only essential where the subject of investigation is itself an aggregate, as a swarm of atoms, or a crowd. But here (mental measurement) the subject is the individual, not the aggregate of individuals as such. This being the case, statistical methods are only necessary in so far as experiment fails to attain its ideal, the ideal of only permitting one causal circumstance to vary at a time. And it should always be the aim of the experimenter not to revel in statistical methods (when he does not revel and swear) but steadily to diminish, by continual improvement of his experimental methods, the necessity for their use and the influence they have on his conclusions. Statistical methods are not only ancillary; they are, to the experimenter, a warning of failure.

“During the last year or two I have made it a practice to begin a short course of lectures to agricultural students, on some elements of statistical methods, by a lecture devoted entirely to showing how exceedingly bad agricultural experiments are, how this introduces the necessity for statistical methods in discussing the results, but does not obviate the primary necessity for improving the experimental methods. Is the latter warning unnecessary in the case of experimental psychology? I can only speak as a statistician with an exceedingly limited and sporadic knowledge of what is being done, but it seems to me that it is urgently necessary. The experiment (if it can be called an experiment at all) is often not only bad, but its badness seems to be accepted merely as a rather interesting fact, calling perhaps for the exercise of more elaborate statistical methods a few partial coefficients of correlation a few rather elaborate probable errors and that is all. It does not seem to be recognized as primarily a condemnation of the experimenter, perhaps because the experimenter has never been exercised in the sciences where really accurate experiment is possible, so that he has no high ideals before him.”

3. Intellectual or emotional repugnance to statistical fatalism has been assigned as one of the reason for objecting to the constancy of the I. Q. I find no reason to believe this to be true of any trained psychologist. The layman and certain schoolmen may

¹Yule, G. Udney. Under, “Critical Notice,” *Br. Jr. Psych.*, vol. XII, Part 1, June, 1921, page 106 ff.

be motivated vaguely by a certain repugnance to anything that undermines the beautiful optimism of,

"I am the captain of my soul,

I am the master of my fate"

but the majority probably do not attempt to think the matter through, or try to orientate intelligence testing in respect to any philosophical basis such as free will or determinism. The practical school man is likely to be an opportunist, and fatalistic finality gives an air of professional expertness, and saves vexatious analysis, trial and error, and suspense.

However, while certainly no psychologist would oppose the constancy of the I. Q. for other than scientific reasons, he is most emphatically opposed to any premature statistical finality regarding the whole matter.

There is nothing gained by dealing with statistical results as causally final or chronologically static. Statistical stability or variability rests upon causes of such a nature that a rational inquiry into the possibility of modification is irksome, undramatic and may become very vexatious by leading to still greater causal complexities. It is also not a simple matter of "either-or" as regards statistical predictability and stability. I read in my morning paper that Los Angeles had over six thousand automobile accidents in 1921. It is safe to predict six thousand for 1922, but it is philosophically bad to resign oneself to the situation. A careful analysis of causes could conceivably reduce the death rate. Likewise statistics could easily have told us that the death rate in building the Panama Canal would have been thus and so, but Goethals did not accept the statistical inference but applied preventive medicine and hygiene.

Now turning to intelligence, it is manifestly a behavior composite. In our carefully non-selected group in the Stanford Revision these is an unbiased sampling of many causal factors of this complex response we call intelligence. Let us name some of the probable ones: heredity, sensory efficiency or deprivation, social milieu, educational opportunities, glandular secretions, general systemic conditions, disease, nutrition, nature of conditioned reflexes, Freudian complexes, emotional make-up, and so on. Immediately it is not likely that any one or a group of these factors will be sufficiently changed, considering the masses, to invalidate the norms of the Stanford Revision. In this sense the age norms are probably temporarily static. It would be bad statistical philosophy, however, acting like an insidious paralysis to all constructive educational advance, if we substituted statistical results for painstaking analysis in a specific case, or for causal analysis for society at large. It is true that there is a wide gulf between

logically possible modifications of causal factors and the surrogation of these factors into the social habituations, which condition gives much greater stability to social action and the complex statistical results than the reformer is likely to appreciate. For example, the means is now probably at hand to wipe out tuberculosis if we had the hygienic habits and hygienic morality to back up the sanitary expert.

Aside from changes in social habituations, unforeseen factors may make marked changes through regulation of glandular secretions. The cure of cretins may be followed by the cure of Mongolian idiots. Again, greater scientific precision in the education of the baby's conditioned reflexes could conceivably change the intelligence of any particular child, and if instituted on a large scale appreciably raise the median performance or change the intelligence rank-order of certain classes of cases.

In pedagogical and administrative problems it is inexcusable to conceive the intelligence norm (namely, the cross-section of the composite set of causal factors) as static. The administrator is justified in making subdivisions in classes on the basis of testing but broad constructive pedagogical and administrative effort must not cease. The teacher should take philosophically the clearly demonstrated cases of mental deficiency and not attempt to make a silk purse out of a sow's ear, to be sure. On the other hand, she should avoid using the I. Q. as a blanket excuse, allowing it to act like a paralysis to all pedagogical resourcefulness and effort. *The low I. Q. of her class can easily become the perpetual alibi of the poor teacher and a new form of infant damnation for the pupils.*

What the average *does* may be pretty safe betting as to what the average will do under a *laissez faire* policy, but it does not follow at all that it is what the average could or should do. Suppose that by scratching Aladdin's lamp we could reduce the enrollment of every teacher to fifteen pupils, or say correct all glandular difficulties or wipe out all superstition, or secure specifics for several of the serious children's diseases, would it not only change the median, P. E.'s, and lower limit but might not whole groups change in their relative position in the rank-order of intelligence? Preventive medicine, refinement of social and sanitary habituations, improved methods of handling the emotions, more individualized teaching, and constructive eugenics constitute our only magic lamp.

Again, sufficient time has not elapsed since the beginning of testing to warrant much prognosis if we are to judge and profit by the vicissitudes of prognosis and cures in medicine.

4. There is considerable evidence from the pathology of the nervous system that the lower end of the intelligence curve marks the

intrusion of a distinct set of factors not found throughout the curve. From normal to superior is probably bridged without any change in kind, from normal to feeble-minded we have a blending of dull normal with varying degrees of pathological derangement; i. e. the entire curve is distinctly bimodal.

This means that preventive medicine or more healthy conditions of conception may remove our worst cases forever. This would not help the pathological cases of to-day to be sure, but such a possibility should make us cautious of statistical estimates of the number of feeble-minded we must have. Equally true, it would be fatuous to feel secure for the future on the basis of the present percentage of low grade intelligence. Relatively changing birth-rates for different classes (all other factors constant) could conceivably shunt the entire intelligence distribution curve. The point is that cogent reasoning and analysis of present and possible factors and not more abuse of the must over-worked normal distribution curve, will solve our difficulties.

5. As pointed out by Stenquist, quite recently, we do not yet know just what a low I. Q. means in terms of eventual adjustment in the world. And as proof of the tester's diagnosis and prognosis lies far ahead in the future (many times the present age of the Binet Tests, Stanford Revision) it becomes peculiarly easy to make sweeping assertions without any immediate factual check.

6. Finally, there is nothing in the nature of the tests to warrant speaking of them as measuring either heredity or environment separately. Any one is, of course, quite within scientific proprieties, when he offers his opinion as to the dominating factor in any mental test performance. We must regret, however, that the commercializing of the tests has resulted in such unqualified statements as this (appearing in the prospectus of a mass test): ".... is an instrument for the measurement in groups of the native ability of any individual whether child or adult, who has had the equivalent of three or four years of schooling."

AN EXPERIMENT IN CLASSIFYING PRIMARY GRADE CHILDREN BY MENTAL AGE.

BY FRANCES LOWELL, PH. D.

DePauw University

Purpose. During the school year beginning September 17th, 1920, the writer, as psychologist for the public schools of Rochester, New York, conducted an experiment in the City Normal School to determine the advisability of classifying children in the first, second and third grades on the basis of mental, rather than chronological, age.

Conditions. The conditions for experimentation were as nearly perfect as one may hope to find in any public school system. Because of the hearty co-operation of the principal, teachers, and school nurse, and because the school used was connected with a City Normal School, the experimenter had absolute freedom in making any desired changes in placing of children, in using all school and health records, and in securing physical examinations of children with presumably remediable defects.

Original Organization of Primary Grades. The general plan of primary organization in the Normal School, as it existed in September, provided each grade, with one exception, with four classes. These classes consisted of two B sections and two A sections. The B sections were beginning first grade work; the A sections had finished one semester. Within each of these B and A groups, a further differentiation was made into average and slow divisions. These were designated by the use of the Arabic numerals 1 and 2 placed after the section letter. Thus, the average First B class would be designated IB1, and the slow First B by IB2; the average First A class would be indicated at IA1, and the slow First A as IA2; the average Second B grade as IIB1, and the slow Second B as IIB2, etc. The average classes were composed of the children who were expected to accomplish the work of their grade satisfactorily, regardless of their chronological ages. The slow groups consisted of children whom the teachers considered too slow in comprehension to succeed in the average group, but too capable mentally to be placed in special classes for subnormals. The exception, mentioned above, to having an average and a slow division for each section of a grade, was in the First A grade, where limited space made it necessary to have just one room accommodating fifty children on half day sessions. The total enrollment in September of the First, Second, and Third Grades was about 300, besides a kindergarten to draw from, which averaged 175 children.

In addition to these average and slow classes, there was a small

group of "problem" cases, known as the "Observation Class" Originally, this group had been composed of about ten children from the first three grades, who either had been absent for several weeks and on their return were found to be in need of special help if they were to continue successfully the work of their grades; or who had some special disability, and after a few weeks in school were found to need more individual attention and help in some subject than the regular teacher could give. These children and others having similar difficulties, were observed carefully and given the necessary aid and encouragement by a teacher trained in ungraded work. Eventually, they returned to their regular grades and others took their places. However, this class had become a "dumping-ground" for all kinds of misfits—mental, social and physical—of the first three grades. In September, it consisted of nine such misfit cases, who were being instructed by a teacher without any special training in the work, and who were making little progress because of the inability of the teacher to deal with the large variety of disabilities exhibited by the different members of the group. At the mid-year reorganization of the grades, to be described later, this class was abolished.

Mental Examination in First Grades. After having carefully surveyed the conditions existing in the first three grades, the writer gave individual mental examinations to the children. The Kuhlmann Revision of the Binet Tests* was used throughout the experiment for obtaining the mental ages of the pupils. The "Observation Class" and all First Grade children were examined before the mid-year promotions occurred. The results of these examinations as obtained from each class, before any changes in grading had been made by the writer, are shown in the following tables.

Table I gives for each grade the average chronological age and the distribution of cases according to chronological age; Table II indicates the average mental age and the distribution of cases by mental age in each grade; and Table III shows the average intelligence quotient and the distribution of intelligence quotients by grades before reorganization.

*Kuhlmann, F.: A Handbook of Mental Tests. 1922.

TABLE I
Distribution by Chronological Ages Before Reorganization.

Grade	No. in Grade	Av. C. A.	5.0-5.4	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	9.5-9.9	10.0-10.4	10.5-10.9
Obs. Cl.	9	8.9				1			2	2	2	1		1
I B1	40	6.8	1	2	13	10	11	2						1
I B1	37	6.7		3	14	11	3	3	2	1				
I B2	18	7.2			1	6	7	2	1	1				
I A1-2	46	7.3			3	9	17	9	4	4				

TABLE II
Distribution by Mental Ages.

Grade	No. in Grade	Av. M. A.	3.5-3.9	4.0-4.4	4.5-4.9	5.0-5.4	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4
Obs. Cl.	9	7.1						2	3	2		1	1	
I B1	40	6.4	1		2		3	18	8	5	1	1	1	
I B1	37	6.4		1		2	4	10	13	6	1			
I B2	18	6.2			1	1	3	8	3		1	1		
I A1-2	46	7.1						8	15	11	5	3	3	1

TABLE III
Distribution by Intelligence Quotients.

Grade	No. in Grade	Av. I. Q.	50-59	60-69	70-79	80-89	90-99	1.00-1.09	1.10-1.19	1.20-1.29	1.30-1.39
Obs. Cl.	9	.80			5	2	2				
I B1	40	.95	1	2	2	5	17	7	3	3	
I B1	37	.96		2	1	6	14	9	4	1	
I B2	18	.87		2	2	6	5	2	1		
I A1-2	46	.98			4	7	16	12	4	2	1

It will be noticed that each table shows two sections of average First B Grade children, i. e., of IB1. This is due to the fact that, in September, the number of children entering the first grade is too large to be accommodated in one room and hence two sections are needed. In the two IB1 groups there were thirteen children who had not yet attained a mental age of six, and nineteen whose intelligence quotients were below .90. Only six children of these groups were below 6 years chronologically, and these were all superior cases.

The slow or IB2 group had five cases with mental ages less than 6 years. Although the slow class was supposed to contain only children with inferior intelligence, or more specifically, dull but not feeble-minded children, this class had eight cases who were normal or superior, and four who were feeble-minded.

The 1A grade showed children ranging in mental age from 6 to 9½ years. While it is generally conceded that a 6½ year mentality is necessary for successful 1A work, eight children in the group were below that age mentally, and eleven had intelligence quotients below .90.

Reorganization of First Grades at Mid-year. By the time all of the first grades had been tested, it was time for the mid-year promotions, and so the writer made the necessary adjustments in placing at that time. The results of these adjustments are summarized in Tables IV-VI. In considering these tables, it must be remembered that the "Observation Class" was abolished, and a class of repeating IB children given to the teacher. Although in January these repeaters all tested over 6 years mentally, and had intelligence quotients above .90, thirteen of them had failed to make their grade because of mental immaturity. Table V shows that these thirteen children, after five months of IB work, had attained a mental age of only 6 to 6½ years. When they entered the grade in September, they must, therefore, have been less than six years mentally. The failures of the other five children in the class were due either to their inability to work successfully in a large group, or to long absences from school. This group of repeaters is designated in Tables IV-VI as IB2. The other IB1 groups of the first semester became 1A1 at the mid-year promotion time, and the slow IB2 became 1A2. The old 1A1-2 section was divided and promoted to IIB1 and IIB2 according as the intelligence quotients of the individuals were above or below .90. From the Kindergarten comes a new group of IB1 and a new one of IB3 children, all of whom were mentally examined by the writer and found to have a mental age of at least 6 years. Those having intelligence quotients above .90 were

placed in IB1, and those with intelligence quotients below .90 were placed in the slow group, namely in IB3.

Table IV gives the average chronological age and the distribution of cases according to chronological age for each grade; Table V presents the average mental age and the distribution by mental ages; and Table VI shows the average intelligence quotient and the distribution of cases by intelligence quotients.

Double Promotions. It is evident, from an examination of Table V, that the re-classification of First Grade pupils was not made, in all cases at least, on a strictly *mental-age* basis. For instance in the newly formed IB1 and IB3 groups, there were eleven children who had a mental age of 6½ years or over, which is the mental age required for successful work in the IA grade. A rigid *mental-age* classification would have allowed these children either to skip the IB class or to do double work therein, i. e., to do the work of the IB and the IA grades in one semester. However, it seemed more advisable to the writer, to accustom them to the routine of school life, after the freedom of the kindergarten, before permitting them to compete with children already habituated—as would be the case if they skipped the IB class—or before allowing them to attempt the extra work involved in double promotions. In the case of the five, mentally over-age, IB2 repeaters, who could not work successfully in large groups, it seemed desirable to place them in a small group of children whose mental ages were slightly lower, in order that their greater ability might aid them in gaining some much needed self-confidence. Hence the same requirements were made of all IB children regardless of whether they were average or superior mentally.

In the IA1 groups there were eleven children whose mental ages would warrant their being placed in a second grade. Instead of being allowed to skip the IA class, these children were given an opportunity to complete the work of the IA and IIB grades in one semester. Since, for efficient handling of groups, each teacher divided her room into three sections, it was a simple matter to put the children who had mental ages over 6½ years into a little group by themselves, and allow them to take the work as rapidly as they were able. To make sure, towards the end of the semester, that the standards of the IIB grade were adhered to, these children were occasionally sent in to share an arithmetic or reading lesson with the regular second grade children. In June, all of these pupils of superior mentality were given double promotions from IA1 to IIA1, since they had completed both IA1 and IIB1 work in one semester. Most of these children after the double promotion, would be working in a grade suited to their mental ability.

TABLE IV

Grade	No. in Grade	Av. C.	A	5.0-5.4	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4
I B1	31	5.8		2	17	8	3	1				
I B2	18	6.2			1	12	4	1				
I B3	16	7.2					3	5		2		
I A1	22	6.7				3	4	10	3	2		
I A1	22	6.6		1		1	8	8	4			
I A2	11	7.2							5	1	2	3
II B1	25	7.3					1	4	10	6	4	
II B2	18	8.3						1	4	7	4	2

TABLE V

Grade	No. in Grade	Av. M. A.	Distribution by Mental Ages.							
			6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	
I B1	31	6.3	21	10						
I B2	18	6.3	13	5						
I B3	16	6.2	15	1						
I A1	22	7.1		15	4	2	1			
I A1	22	6.6		18	4					
I A2	11	6.6		10	1					
II B1	25	7.6			12	8	2	1		
II B2	18	7.0			3	12	3			

TABLE VI

Distribution by Intelligence Quotients.										
Grade	No. in Grade	Av. I.	Q. 70-79	80-89	90-99	100-109	110-119	120-129	130-139	
I B1	31	1.07			6	14	10	1		
I B2	18	.99			12	6				
I B3	16	.86		16						
I A1	22	1.07			8	6	5	3		
I A1	22	1.00			12	6	3	1		
I A2	11	.84	1	10						
II B1	25	1.04			9	11	2	2	1	
II B2	18	.84	2	16						

The same conditions existed in the IIB grade, and the superior children were given double promotions in June from the IIB1 to the IIIB1, as they had completed the work of both the IIB1 and the IIA1 grades in the one semester.

Although, as stated above, the writer employed a double promotion method, and regards it as the method, under the existing conditions, best suited to the primary grades to bring about adaptation between school grades and the mental ability of children, she does not wish to be regarded as giving it an unqualified endorsement. Double promotions should probably not be continued beyond the fourth grade, because such a course would result in children finishing the grades long before they were physically or socially ready to mingle with high school students.

Demotions. In the IB grade, twelve children were found whose mental ages, at the mid-year, were less than 6 years. These were demoted to the kindergarten, though some of them had repeated IB work two and even three times. The joy of finding something in school that they could actually do, compensated for their failure to complete the IB work at that time. The parents of only two of these children inquired about the demotion, and to these the writer explained that the children had become so discouraged over their inability to do what the other pupils were doing that they had ceased even to make an effort; that while they were chronologically old enough for first grade, their minds were still too immature for it; that by placing them in the kindergarten for a little while longer, where they could do the work, and in fact, be leaders, their self-confidence would be restored to such an extent that when they later were mentally ready for the first grade, they could do the work and do it well.

The writer is convinced that when children have spent the allotted two years in the kindergarten, and still are not ready mentally to undertake first grade work, provision should be made for a "pre-primer" class. This should be a sort of preparatory class in which work more advanced than that of the kindergarten, but less difficult than that of the first grade, should be given. Much physical training and language work might well be the basis for a course of this kind. The need for a "pre-primer" class is strongly felt in cases where children are overdeveloped physically but underdeveloped mentally, and yet are too young to be placed in a special class for the feeble-minded.

Mental Examinations in Second and Third Grades. During the second semester, from January to June, the writer gave individual mental examinations to the children in the remaining

TABLE VII
Distribution by Chronological Ages Before Reorganization.

Grade	No. in Grade	Av. C. A.	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	9.5-9.9	10.0-10.4	10.5-10.9	11.0-11.4
II A1	37	8.0	3	5	11	10	3	3	2			
II A2	26	8.4			5	10	5	4		2		
III B2	24	9.5				4	5	1	5	8		1

TABLE VIII
Distribution by Mental Ages.

Grade	No. in Grade	Av. M. A.	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	9.5-9.9
II A1	37	8.5		2	7	7	10	9	2
II A2	26	8.1	1	2	8	10	3	1	1
III B2	24	8.5			4	8	6	4	2

TABLE IX
Distribution by Intelligence Quotients.

Grade	No. in Grade	Av. I. Q.	70-79	80-89	90-99	1.00-1.09	1.10-1.19	1.20-1.29	1.30-1.39
II A1	37	1.06	1	4	3	15	8	4	2
II A2	26	.96	3	5	9	5	3	1	
III B2	24	.89	3	10	8	2	1		

sections of the second grade, i. e., to the IIA1 and the IIA2 and to the slow section of the IIIB grade.

Table VII gives for each grade the average chronological age, and the distribution of cases according to chronological age. Table VIII shows the average mental age and the distribution by mental age for each grade; and Table IX gives the average intelligence quotient and the distribution by intelligence quotients per grade.

Inspection of Table VIII shows that twenty-eight of the thirty-seven children in the IIA1 grade had a mental age of 8 years or over, which is the mental age required of *third* grade children. Even in the slow IIA, more than one-half of the pupils possessed third grade mentality. In the slow IIIB class, six children were mentally ready for the fourth grade. The distribution of cases by intelligence quotients is still more interesting. Table IX shows that in both the IIA1 and the IIA2 groups, the range of quotients is from .70 to 1.40 or, in terms of brightness, from feeble-mindedness to "near genius"; while in the slow IIIB group, it is from .70 to 1.20, or from feeble-mindedness to a high degree of superiority.

It was nearing the end of May when these examinations were completed, so the changes in classification were not made until the time for June promotions. Tables X - XII represent the organization of these groups as it existed at the end of the school year, after a classification by mental age had been made by the writer. They show, for each grade, the distribution of cases according to chronological age, mental age and intelligence quotients, respectively.

Reorganization of Second and Third Grades in June. In reorganizing these groups, mental age again formed the basis for selection, and where possible, brightness also was considered. In selecting children for the two sections of IIIB1, it was necessary to remember that the purposes of the groups were different. The first section was to be composed of children whose mental ages were, or would be by September, sufficiently high, and whose intelligence quotients indicated brightness of such a degree, that the children would be capable of doing double work, that is, the work of both the IIIB and the IIIA grades, in one semester. The second section of the IIIB1 would do only the regular IIIB work during the term. With this in mind, the writer selected for the first section thirty-three children whose average mental age was 9.1 years, and whose average intelligence quotient was 1.11. It will be noted, however, that several superior children were placed in the second section of the IIIB1, while

TABLE X
Distribution by Chronological Ages After Reorganization.

Grade	No. in Grade	Av. C. A.	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	9.5-9.9	10.0-10.4	10.5-10.9
III B1	33	8.3	3	5	16	7	2			
III B1	18	8.3		5	5	6	2			
III B2	16	9.5				3	5	5	2	1
III A2	18	10.0				2	3	3	4	6

TABLE XI
Distribution by Mental Ages.

Grade	No. in Grade	Av. M. A.	7.5-7.9	8.0-8.4	8.5-8.9	9.0-9.4	9.5-9.9	10.0-10.4	10.5-10.9
III B1	33	9.1		4	9	9	8	2	1
III B1	18	8.6		9	7	2			
III B2	16	8.2	1	12	3				
III A2	18	8.6	1	8	3	6			

TABLE XII
Distribution by Intelligence Quotients.

Grade	No. in Grade	Av. I. Q.	70-79	80-89	90-99	1.00-1.09	1.10-1.19	1.20-1.29	1.30-1.39
III B1	33	1.11			2	13	12	4	2
III B1	18	1.04				5	10	2	
III B2	16	.87		2	9	5			
III A2	18	.87	1	1	11	6			

several average children were included in the first section. This was due to the fact that the superior children mentioned had been absent for several weeks with whooping cough; and while it was thought probable that they could overcome this handicap sufficiently to be given a single promotion, it did not seem advisable to allow them to be doubly promoted. Consequently, the first section of the IIIB1 had to be completed by a few children of average ability who possessed, in addition to the mental age necessary for the work, excellent health, remarkable persistence and fine habits of study.

Further consideration of Tables X - XII brings out certain facts regarding the two slow groups IIIB2 and IIIA2. While the difference in the average mental age of these two groups is approximately one-half year, yet the distribution of cases according to mental age shows but slight variation in the two groups. In fact, half of the IIIA2 pupils possess mental ages which, under ordinary conditions, would indicate that they belong in the IIIB2 group, i. e., they possess 8 year instead of $8\frac{1}{2}$ year mental ability. Why, then, it will be asked, were they placed in the IIIA2 class along with children who really possess third A mentality? Because they had already repeated the IIIB2 grade twice! Advancing these children to a higher grade with its new interests was the only existing alternative to forcing them, for a fourth time, through the routine of the hated IIIB2.

Although many school systems are providing special courses of study for their feeble-minded and superior children, few have yet arranged a special curriculum for their slow, dull children. In the first and second grades, it is largely a matter of time before the dull child can accomplish, in some fashion, the work required of the average pupil in the same grade, but in the higher grades, it becomes practically impossible for him to master all phases of the work, even when the time is unlimited. He keeps repeating grades until he has not only lost all interest in the school work, and ceased to make any effort to learn, but until he has lost his self-confidence, and indeed, even his self-respect. Until schools provide a special curriculum adapted to the needs and limitations of these dull children—children with intelligence quotients from .75 to .89 inclusive—little is to be gained from classifying them on the basis of mental age. And yet, 34%, or more than one third of all the primary grade children in the school used in this study, and it was a representative school, belong to the dull group.

What is the significance of an intelligence quotient of .75? Or more specifically, what is the outlook for such children as those in the IIIA2, whose chronological age is 10.5 years, whose mental

age is 7.9 years, and whose intelligence quotient, therefore, is .75? In their 4½ years of school attendance, assuming that they entered when they were chronologically six, they have completed 2½ years' work. If they continue their school life until they are 16 years old, at which age the law, in most states, permits them to leave school and go to work, they will have had 5½ years more of school attendance, which will presumably place them in the VIA grade. At that time their mental age would be 12 years, and the intelligence quotient, of course, would still be about .75. What are these children fitted to do when they leave school at the age of 16 years? Nothing! They have been struggling with arithmetic, history, geography or grammar, instead of acquiring knowledge of practical value. Still, school systems do not feel it incumbent upon them to provide a curriculum which will fit the dull child to earn his living and cope with his environment.

The changes made in reorganizing the IIB2 and IIIA2 grades were of two kinds: First, some normal children, who had been misplaced in the IIB2, were removed to average groups; and second, the older and more capable children in the IIIA2 were placed with the repeaters from the IIB2. Time was too brief to permit any examination to be made in the IIIA1 grade.

Mental Examination of Kindergarten Pupils. During the last month of school, all kindergarten teachers, in the Normal School, reported the names of the kindergarten children who would be ready, chronologically, for entrance into the first grade, in September. To these, the writer gave individual mental examinations. The group totalled fifty-seven cases which were disposed of as follows:

- 44 - Promoted to IB1 or IB2.
- 1 - Sent to Institutions for Feeble-minded.
- 1 - Sent to Primary Special Class for Subnormals.
- 11 - Retained in Kindergarten because they would not be mentally 6 years by September.

The results of the examinations of the forty-four kindergarten children who were promoted in June, 1921, to the IB1 and IB2 grades are presented in Tables XIII - XV. Table XIII shows the distribution of cases by chronological age; Table XIV gives the distribution by mental age; and Table XV by intelligence quotients.

Effect of Mental-Age Classification on Promotions. A careful study of data, obtained from an experiment such as the one discussed in this paper, brings forth a very practical question. If children are graded according to mental age, will retardation in schools be greatly reduced? In other words, will the percentage

TABLE XIII
Distribution by Chronological Ages.

Grade	No. in Grade	Av. C.	A.	5.0-5.4	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	7.5-7.9	8.0-8.4
I B1	32	5.9	3	16	11	1	1	1		
I B2	12	6.9			2	4	4	1	1	1

TABLE XIV
Distribution by Mental Ages.

Grade	No. in Grade	Av. M. A.	5.5-5.9	6.0-6.4	6.5-6.9	7.0-7.4	
I B1	32	6.4	2	14	14	2	
I B2	12	6.1	4	7	1		

TABLE XV
Distribution by Intelligence Quotients.

Grade	No. in Grade	Av. I. Q.	70-79	80-89	90-99	1.00-1.09	1.10-1.19	1.20-1.29
I B1	32	1.09			3	16	10	3
I B2	12	.88	1	5	6			

of honestly earned promotions be nearly 100? A rather convincing answer to that question is given by the records showing the percentage of promotions for the first grades of the City Normal School before the classification was made and also afterwards. These figures represent both slow and average groups combined.

June 1920 — 62%

June 1921 — 87%

If we analyze the percentage of promotions for June, 1921, into its component parts we find a still more convincing reply to our question. Considering just the IB1 and IA1 groups, or 66 cases, the percentage of promotions was 97, and 23% of these were *double promotions*. Considering just the slow groups, or 41 cases, only 68% were promoted. Likewise in the IIB1 and IIA1, or 64 cases, there were 100% promoted. Yet in the slow second grades, i. e., the IIB2 and IIA2, there were only 75% promoted.

The above data indicate, then, that the average pupil, when placed according to mental age in a grade, will successfully accomplish the work of that grade, since the curriculum is fairly well adapted to the ability of the average child. The dull child, on the other hand, even when placed according to mental age in a grade, will be unable to accomplish the work in the time allotted, because of his slower rate of mental growth. Consequently, when no special curriculum is provided for him, he will have to repeat the grade, and the bad habits of study, the loss of interest and of self-respect, the attitude of discouragement and hopelessness which accompany the repetition of a grade, continue to cause further failure, even though at the time of the second trial, there is sufficient mental ability to insure success.

Summary. From the data obtained from this experiment in classifying primary children on the basis of mental age, the following conclusions have been suggested:

1. In grades where classification is not based upon mental age, one finds between the children of a single class as great a range in mental age as $5\frac{3}{8}$ years, and in intelligence quotients as great a difference as 79 points, yet these children are all trying to accomplish the same tasks and assimilate the same amount of subject matter.

2. Both mental ages and intelligence quotients must be considered in arranging classes. Children may have the same mental ages and yet differ so greatly in chronological ages that it would be impracticable for them to try to work together.

3. When mental age and brightness are used as a basis for classification, it is possible to separate the dull and feeble-minded children from the average and superior individuals, so that each may com-

pete with others of his own mental calibre.

4. Slow groups should be formed from the first grade up, and a special curriculum worked out which will fit these children to earn a living, since most of them never get beyond the sixth grade.

5. The superior children, instead of skipping grades, may be given double work, which will bring them up, in a semester or two, to the grade for which they are mentally ready, without the loss of any fundamental part of the school work.

6. Where first grade work is standardized upon the basis of the performance of average children 6 years old chronologically, no child should be allowed to enter the first grade before he has attained a mental age of 6 years. If he is chronologically six, and has been in the kindergarten two years, or if the school system does not provide for kindergarten training, he should be placed in a pre-primer or "preparatory-to-first-grade" class until he becomes 6 years mentally.

7. When the teaching and the curriculum are adequate, and when children are not exposed to the work of any grade until they are mentally qualified for it, a mental age classification will give practically 100% of promotions in the primary grades.

SPEECH AND WILL-TEMPERAMENT

LOVISA C. WAGONER AND JUNE E. DOWNEY

University of Wyoming.

So great is the present-day interest in speech defects that no apology is needed for attempting to develop a new line of attack.

Estimates of the percentage of the school population who are in need of remedial treatment for the various sorts of speech defect vary from three to five and one half per cent. The social and vocational handicap of defective speech is obvious. In the last few years clinics and courses for training in the use of corrective measures have multiplied. Interest has grown by leaps and bounds and has been stimulated by work done with speech defects resulting from injury during the war. This widespread concern has spent itself very largely in attack upon symptoms. Experimental investigation has been exceedingly limited.

We propose to attack the problem of speech from two different angles.

1. Do different varieties of speech defectives, when given the Downey Will-Temperament Test,¹ show any characteristic profiles? Investigation of the temperamental and volitional qualities which underlie individual variations in timbre, volume, and rate of speaking and the changes which are commonly recognized as important tokens of mood and attitude may throw light upon both the normal and the atypical factors in speech.

2. Is the graphic will-temperament test, in which writing exercises are used to determine the dynamic or activity level of the individual, translatable into a vocal test in which vocal exercises are used? If so, do the two profiles obtained by these different methods of procedure resemble one another in any way? The answer to this question bears upon the general validity of the will-profile and may point the way to a profitable study of speech difficulties both in their mild and in their aggravated form.

Before attempting tentative answers to our questions, let us call attention to a number of similar leads in the literature on speech defects.

Scripture stresses importance of the voice as an indication of personality for he says, "It is unquestionably true that mental or bodily weakness show themselves clearly in the speech of the individual." Again, "An excess of energy in articulation often occurs. This may be due to increased vitality of the organism whereby all movements are affected or to diminished accuracy of control."²

¹See "Manual of Directions, Downey Individual Will-Temperament Test," World Book Company.

²"Elements of Experimental Phonetics," 1902, Scribner, p. 270.

Notes.

The slovenly speech of the mental defective is commonly credited to his poor physical and mental tone. Tredgold, in speaking of the speech defects of the feeble-minded, says, "The chief defects consist of a thickness and indistinctness of utterance, an imperfect articulation of consonants, and (rarely) stammering and stuttering. The former of these conditions is partly attributable to abnormal configuration of the palate, lips, jaws, or pharynx, and partly to a general *brain inertia or unwillingness to make the necessary effort for distinct enunciation*. The consonantal defects are due to *similar causes plus a want of coordination*." (italics ours). This inability properly to coordinate impulses accounts for the tediousness and the lack of permanent results of corrective speech work with such individuals. The fact that imperfect or negligent speech is not a constant characteristic of mental defect of a given degree, indicates that there is something besides general intelligence involved. To state the cause in such terms as "faulty coordination of impulses" is not an adequate explanation except in cases showing paralysis. It is necessary that a more penetrating analysis be made not only of the untidy speech of defectives but of the negligent vocal habits of normal children and adults as well. There has been little tendency to search for underlying reasons which would explain faulty pronunciation. Only two remedies have been generally applied, correction of anatomical and physiological defects, and formation of new habits.

Stuttering has long been attributed to emotional disturbances such as dependence on parents, a feeling of inferiority, and repression, but the temperamental traits which provide rich soil for such unwholesome growths have been very largely ignored. The corrector of speech defects tacitly recognizes the fundamental nature of the temperamental bias underlying the symptom we call stuttering in that he stresses the importance of stimulating self-confidence and of reducing self-consciousness to a minimum. This he does largely through transfer to the patient of his own buoyant attitude and zeal. His trial and error method sometimes results in accidental success which is permanent in proportion as he has been fortunate in using methods suited to the temperamental and volitional constitution of the particular individual with whom he is working. His lucky strike may be hailed as a "cure."

As early as the first half of the nineteenth century Marshall-Hall and Lichtinger believed that there exists as a basal fact in the etiology of stuttering insufficient "impulses of volition." Despite the prevailing tendency to find an anatomical basis with

³Opus Cit. p. 466.

⁴"Mental Deficiency," 1921. Wm. Wood and Co. New York. p. 168.

the resulting frequency of surgical measures for relief, Merkel⁵, who was a stutterer, believed the cause to be altogether psychic and to lie in "the sphere of the will." In 1911, Appelt⁶ brought out his most valuable book in which he attributes stammering to a disturbance occurring in childhood, the results of which may be removed by Freudian measures. It is difficult to accept this theory as adequate explanation of all cases.

Some experimental work has been done along physiological lines, such as that of Robbins' with the plethysmograph in measuring increases in brain volume in a trephined stammerer. "Stammering was accompanied by much more marked increase in brain volume than could be accounted for by either the physical or mental work used in normal speech." "It is reasonable to conclude that increase in brain volume is an important factor in the production of stammering." Robbins' work was undertaken with especial reference to Bleumel's theory that stammering is due to cerebral congestion. It may well be that the tone of the vasomotor system bears a close relation to the conflict between temperamental traits which is discussed below.

A most interesting attempt to secure reliable data which will prove or disprove the validity of certain assigned causes of stuttering is reported by Anderson from the laboratory of the University of Wisconsin. "The purpose is to present briefly some results of a number of special tests designed to detect and measure certain traits in reactions outside of the field of speech, and given in such a way as to show, if possible, whether these traits are definitely associated with stuttering. In view of the large number of alleged causes, this experimental work was planned so as to constitute a survey of significant types of mental and physiological processes which are common to vocal and non-vocal functions." The tests were aimed principally at innate "or at least internal" factors. There were three groups or types of subjects, persons of normal speech, those who stutter, and ex-stutterers. The performance of these three groups was compared. The results of a foot-tapping test which was given to reveal a general defect of physical coordination, should such exist, showed no characteristic difference between stutterers and normals but much greater regularity and greater rapidity on the part of ex-stutterers. In order to test rhythm and coordination in relations other than speech, the subject was put through a "complicated series of arm movements, in some trials

⁵"Physiologie der menschlichen Sprache." 1866.

⁶"Stammering." 1911 Methuen.

⁷"A Plethysmographic Study of Shock and Stammering in a Trephined Stammerer." *Am. J. Physiology*, (1919) 48: 286-330; (1920) 52: 11-24.

⁸"Preliminary Report of an Experimental Analysis of Causes of Stuttering." *J. Applied Psychol.* (1921) 5: 340-349.

following the beating of a metronome, and in other trials simply going as fast as he could." The results indicate that stutterers "tend to do relatively better when following the metronome" and thus conform to the generally accepted notion regarding the effect of rhythm on the stutterer's speech.

In memory span for movement stutterers are markedly inferior to normals; the inability "to hold several things in mind at the same time" is thrown into relief by a block test. The stutterer seems to "have difficulty in grasping several movements simultaneously. The significance of the kinaesthetic factor seems to be that in consciously directed speech movements, a great many kinaesthetic elements must be held in mind and thus coordinated."

Referring to the moot point as to whether a particular form of imagery is at fault the writer gives the opinion that the difficulty is not with imagery itself but rather with the coordination of imagery. "The results from the visual imagery tests show more relation to intelligence than to severity of stuttering."

One of the most suggestive remarks is with reference to inhibition. In some stutterers there is a correspondence between lack of inhibition in the test, which consisted of responding "by raising his arm slightly every time 'three' was read immediately after an odd number but not after an even number," and lack of inhibition in speech and other behavior. "Stutterers differ as to whether they will or will not start a wrong response but tend to agree in inability to check or modify the response after it has been started. His response tends to be of the all-or-none type." In short "it seems very probable that a general lack of ability to inhibit an impulse after it has found partial expression is an essential factor in stuttering."

This work of Anderson (it is to be hoped that a more complete account will be published) is important not only on account of its experimental attack but also because of its suggestion that there are innate traits and combinations of traits which are conducive to the development of speech disorders. Perhaps it would state the situation more adequately to say that speech defects are the outward manifestation of temperamental and volitional traits and that the only hope for "cure" which will be genuine and not merely apparent lies in insight into traits of personality. If this is true the practical value of a temperamental scale for the study of speech disorders is obvious. Complexes, blocking, and repression are only symptoms. The real problem to solve is why the complex or the blocking exists.

Taking as our thesis the proposition that the fundamental factor in speech difficulties frequently lies in the temperamental make-

up of the individual, the writers plotted the will-profile of those individuals who by reason of a speech difficulty had been referred to one of us for corrective treatment.* Our cases are not numerous and our conclusions are only tentative but they do suggest a positive answer to our first question, namely the possession by the speech defective of a will-profile that suggests the reason for his difficulty.

Our records indicate several possibilities which the material at hand is as yet insufficient to prove.

The individual who speaks with great deliberation or hesitancy or, in more extreme cases, drawls, gives a will-temperament profile running low on speed of movement and of decision, low on motor impulsion and finality of judgment and high on interest in detail. If, in addition, such an individual is working at his best speed (the profile shows freedom from load) he may, under external pressure, actually block.

It is, of course, not possible to assert from the material at hand that all persons giving the will-profile outlined above would be hesitant in speech. But extensive experience with will-temperament testing has shown that it is quite possible to select subjects for a "deliberate" profile by a preliminary observation of speech characteristics.

Conflicts in temperamental traits might well cause speech blocking, which in extreme cases would lead to stammering or stuttering. Our profiles of occasional stammerers show considerable variety but frequently suggest possibilities of conflict, for example, high speed of decision with low speed of movement and low motor impulsion. The common features seem to be low motor coordination, low motor impulsion, and low motor inhibition. In these respects our results agree with Anderson's. It is certainly noticeable that in our group of students showing mild speech defects, none score beyond the median for motor impulsion, while a group of college debaters all score above the median. Public speakers whom we have tested run very high on motor impulsion.

The graphic test for motor impulsion, it should be recalled, consists in the maintenance of handwriting under various distractions. Under such conditions the writing of the hyperkinetic or explosive individual becomes larger and speedier as his writing movements become more and more automatic. For the inhibited, blocked individual, under the same conditions, writing shows signs of tension. It becomes reduced in size, heavy and slow. All gradations occur from the excessive magnification and speed evident in the very

*The experimental work for this report has been done by Miss Wagoner, who has had special training in the theory of speech defects and much practice in corrective speech work. J. E. D.

explosive individual to the tiny crawling hand of the excessively inhibited.

Something very similar is noted in speakers. The orator as he gets interested in the situation and "lets himself go" shows more and more freedom in voice and gesture. Both volume of voice and speed of vocalization may show increase. The speaker who lacks such power of expression as he withdraws his attention from his voice begins to speak hesitantly in a thin reduced head voice. He can "put over" beautifully a memorized and thoroughly practiced speech but cannot trust himself to do big things in an impromptu way.

There is, of course, a third possibility. Too great motor impulsion and speed with deficient coordination and motor control (motor inhibition) may give the speaker who "stumbles over himself" whose drive and speed result in chaotic presentation; or the speaker of the piercing voice or disagreeable nasality of tone. His attention is completely diverted from the auditory side of the situation. Careful training and the acquisition of proper habits are needed to safeguard his "oratory" from his own enthusiasm.

There is a type of speech-defect called "cluttering" with many varieties from negligent to slovenly speech. It appears to us—and such records as we have bear it out—that the clutterer is a speedy inertialess individual of high motor impulsion but lack of balance on the side of control. These "cluttering" records are most common in the profiles of adolescents, just where one might expect to find them.

Our observation of the voice changes taking place under varying conditions led us to attempt a translation of the graphic will-temperament test into a vocal test. A little experimentation indicated that this could be done with comparative ease. Our interest in this connection was not only to get a tool for analysis of speech defects but, as stated earlier in the paper, to test further the validity of the graphic will-temperament test, for it is evident that any degree of correspondence between two profiles obtained by the use of different activities would be a strong argument in favor of the significance of the original series.

If only minor discrepancies are found for will-profiles obtained from a graphic and a vocal type of activity an opportunity is given to study the relationship between temperament and the formation of specific habits. It is even conceivable that a closer parallelism might be found to exist between graphic and vocal profiles for normal individuals than for speech defectives. Such a result would support some interesting hypotheses concerning speech disorders.

In paralleling the graphic test the first consideration was the

selection of a vocal activity which would give opportunity for play of pitch, volume, and speed but would not depend upon memory. Because in adults the repetition of the alphabet is reduced to an automatic response, the vocal test employs repetition of the alphabet under various conditions as its medium. First the natural speed of vocalization and the normal pitch and volume are measured and then compared with the results of speeded recitation. Flexibility is rated in terms of the ability to shift pitch and volume. Precision of enunciation in reading test phrases would seem to give indication of care for detail. Coordination of impulses is measured by the speed of vocalization with a double task in view; rapid recitation with plastic lips, for example. The length of time required to re-adjust after interruption together with the ratio of normal to speeded recitation provides a tentative index of the freedom from load. Ability to maintain rate and quality of vocalization while maintaining other activities such as reading silently, indicates degree of motor impulsiveness. Motor inhibition is expressed by the degree of retardation in recitation. Reaction to contradiction is tested by the response to the statement that a given letter has been omitted in recitation—an auditory test in place of the visual one of the graphic series. Finally, emotional stability is measured in terms of vocal reaction when a shrill whistle is blown during the recitation.

As norms have not yet been established it is possible to give only an estimate of the correspondence between the two forms of the will temperament test in terms of coefficients of correlation.

Twenty-three subjects were ranked according to their scores on the graphic and vocal form of the will-temperament test for the nine traits for which parallel tests had been determined. The correlation coefficients by the rank method were positive and high for motor impulsiveness (rated on time only) motor inhibition and reaction to contradiction; positive and mediumly high for flexibility, and interest in detail (scored in vocal test on flat number of errors); and positive but low on coordination of impulses and volitional perseverance. For freedom from load and speed of movement there was no indication of relationship. In the graphic test the scored records are preceded by preliminary practice in order to overcome self-consciousness and it is proposed to give such preliminary practice in the vocal test and determine the effect of this upon the correlation for speed.

On the whole the preliminary experiments with the vocal test have been very encouraging in their outcome. Revision on the basis of this preliminary work and establishment of norms is now in progress.

THE CONVERSION OF TEST SCORES INTO SERIES WHICH SHALL HAVE ANY ASSIGNED MEAN AND DEGREE OF DISPERSION.

BY CLARK L. HULL
University of Wisconsin

It is frequently desired to convert the test scores from a group of subjects into a new series, which shall be directly comparable with some other standard series. It may be desired, for example, to convert the scores from an educational test directly into school marks. More frequently, it is desired to convert the scores from a number of mental tests into strictly comparable series for the purpose of plotting mental profiles. It is obvious that in order to be comparable, the respective series of scores must have the same mean and the same degree of dispersion as measured by the mean variation or the standard deviation.

Such conversions may be made very readily and accurately by the following method. In case a large test score is to correspond to a large score in the converted series, the formula is:¹

$$X = K + SX_1$$

where

M is the mean of the converted series

σ is the S. D. of the converted series.

X is any individual score of the converted series

M_1, σ_1, X_1 , are the corresponding values of the test series

and where

$$S = \frac{\sigma}{\sigma_1}$$

$$K = M - SM_1$$

The method may be illustrated by a solution based on the following data:

$$M_1 = 30$$

$$M = 80$$

$$\sigma_1 = 3$$

$$\sigma = 6$$

$$X_1 = 27$$

X is to be found

Accordingly,

$$S = \frac{6}{3} = 2$$

$$K = 80 - 2 \times 30 = 20$$

Substituting in the formula,

$$X = 20 + 2 \times 27$$

$$X = 74$$

¹The mathematical derivation of this formula is not difficult and may easily be supplied by the interested reader.

It will be observed that S and K need be computed but once in any given series of conversions. This being done, the conversion itself is a very simple matter.

Where the series to be converted involves any considerable number of scores, it will usually be advisable to construct a table. With the aid of a good computing machine, such as the Monroe, this may be done in five or ten minutes.¹ K is placed on the lower dial, S on the keyboard and the crank is turned once for every value of X_i , a reading (X) being taken at every revolution until the range of X_i has been traversed. Such a table may be preserved for the conversion of subsequent scores of a comparable nature which may be secured.

It sometimes happens that a *small* score in the original series corresponds to a large score in the converted series. In this case the formula becomes:

$$X = K - SX_i$$

where all values are the same as before, except that

$$K = M + SM_i$$

In either formula the S may be computed from either the two standard deviations as indicated above, or from the two mean variations as desired.

The conversion of test scores into special series for the plotting of mental profiles and other detailed comparisons, raises the question as to what are the most desirable characteristics of series for such purposes. One of the most obvious is that the series should be in a form, if possible, which is already widely used and appreciated. Such a series is found in the ordinary marking system. Test scores may be translated into school marks by giving M the value of 81.5 and σ the value of 6.

Some years ago Woodworth proposed for this purpose the use of an index which was 100 times the ratio of the deviation of the individual test score from the group average, to the standard deviation of the scores of the group.² This may be readily obtained by the above method if M be given the value of zero and σ the value of 100. This method yields a series of plus and minus values ranging from zero to somewhat above 300 in either direction. Besides involving rather large and clumsy numbers this method suffers from the fact that it is probably intelligence to few except professional psychologists.

Perhaps the most promising system is one where the converted scores will range from zero to 100 as practical limits with the mean at 50. Since individual deviations from the mean test score

¹In most cases an ordinary adding machine may be used in the construction of such a table.

²Psychological Review, 1912. Vol. XIX, p. 97.

will rarely exceed 3.5 times the standard deviation, the sigma of this system will accordingly be 50 divided by 3.5 or approximately 14. As a final illustration of the method, two short series of scores will be converted according to the system just mentioned. The original scores are given in column A and B of the following table:

	A	B	A'	B'
	2	18	32.5	32.5
	10	72	60.5	53.5
	6	54	46.5	46.5
	13	126	71.0	74.5
	4	45	39.5	43.0
	<hr/>			
Mean	7	63	50	50
σ		36	14	14

The S and K having been computed in each case, the formula for series A becomes,

$$X = 25.5 + 3.5X_s$$

and that for series B becomes,

$$X = 25.5 + .389X_s$$

Substituting the values of series A and B in their respective formulae, the converted series A' and B' are obtained. It will be noted that both converted series alike have a mean of 50 and a standard deviation of 14 as would be expected.

NOTES AND NEWS

ANNOUNCEMENT.

With the June issue the ownership of the Journal of Applied Psychology passed from Clark University to the present and former Editors, James P. Porter of Ohio University, and William F. Book of Indiana University. The delay in getting out the present issue is due to the additional labors involved in this transfer and to the change that was necessary in the place of its publication. It is the desire of the editors not only to issue the Journal promptly in the future, but to improve it in certain important respects and if at all possible without extra expense to subscribers. The subscription price of the Journal has remained the same throughout the entire war and reconstruction periods. It is the intention as conditions in the printing world settle down to normal to strengthen the periodical in every possible way. Since the former editors will remain in charge and assume full responsibility for its success, no fundamental change in policy will be made. A more definite statement in regard to future purposes and plans will be announced in the December issue.

The committee on Information and legislation of the American Psychological Association, section Clinical Psychology, has published its report upon "Existing laws which authorize psychologists to perform professional services." (*Journal of Criminal Law and Criminology*, Vol. XIII., No. 1, May, 1922. pp. 70-73.) The report shows that laws mentioning the services of psychologists exist in Illinois, California, Kansas, Oregon, Wisconsin, New York, South Dakota, Tasmania. These laws have all been passed since 1915. In most of the cases the psychological services referred to are that of judging, regarding insanity, feeble-mindedness, moral delinquency, before commitment.

At Stanford University Dr. Lewis M. Terman has been appointed head of the department of psychology to succeed Professor Frank Angell who retires at the end of the academic year 1921-1922. Other additions to the department include Dr. W. R. Miles, professor, and Dr. Calvin P. Stone, assistant professor.

A NOTE ON THE EXTENT TO WHICH SYSTEMS OF
CHARACTER ANALYSIS ARE USED IN THE
BUSINESS WORLD.

Workers in the field of applied psychology are frequently called upon to express opinions of so-called character analysis systems. In fact, these systems (especially that of Katharine Blackford) are so widely referred to that a person in contact with managers of sales and personnel is likely to over-estimate the extent to which the systems are used. Psychologists not in close touch with business, on the other hand, are even more likely grossly to under-estimate the part these pseudo-scientific methods play in the business world. No data, so far as we are aware, are available as a basis for estimating the actual extent of the use of character analysis.

A limited questionnaire study was undertaken in order to gain some answer to this question. Two hundred questionnaires were mailed, one hundred to employment managers of industrial plants employing over three hundred workers, and one hundred to insurance agency managers. The companies to which the questionnaire was sent were selected at random within these groups. All were in the city of Chicago. The principle questions on the blank were: "Do you make use of a system of character analysis in sizing up men?" and "What is the name of the system used?" (An introductory note explained what we meant by systems of character analysis.) Other questions sought to obtain the individual's opinion of these systems. The position, age, and schooling, of the men reporting were also secured.

Sixty-five replies were received, twenty-two from the insurance companies and forty-three from the industrial plants. Six of the replies reported the use of some system of character analysis. Four of these were from employment managers and two were from agency managers in insurance companies. Three of the six used the Blackford system two used a combination of Blackford's with some other system; one did not specify the system used. Five of the six men reporting the use of character analysis systems are university graduates and the sixth is a high school graduate.

Six in two hundred is probably a closer estimate of the use of character analysis within the groups investigated than is six in sixty-five, since persons using such methods would be much more likely to reply than would those acquainted with the systems mentioned in the questionnaire.

A. W. Kornhauser

University of Chicago

A. W. Jackson

BOOK REVIEWS

SANTE NACCARATI, *The Morphologic Aspect of Intelligence.*

Columbia Contributions to Philosophy and Psychology,

Vol. XXVII, No. 2. New York, 1921, 44 pp. Price \$1.10

This is an attempt to determine the extent to which certain anthropometric traits correspond with intelligence. Previous investigations in this field have shown discordant results, the reason ascribed by the author being that the anthropometric traits considered were used singly. Intelligence, however, is a complex trait; and accordingly any anthropometric trait that is to be considered as a correlative must be compound, made up of several elementary traits.

To secure this the author devised a "morphologic index" given by the ratio of value of the extremities (length of one upper and one lower limb) to the volume of the trunk. A preliminary investigation showed a correlation of $+0.75$ between ratio of height to weight and ratio of limb to trunk value. In the light of this high correlation the author assumed the ratio of height to weight as an approximate indicator of the morphologic index, and took that ratio as a simpler method of study for several groups whose morphologic measurements could not be secured.

The subjects used were University students who had been measured with either the Army Alpha, Otis, or Thorndike intelligence tests. The average coefficient of correlation between intelligence and the ratio of height to weight among 221 subjects was $+0.22$. A group of 75 male subjects showing a correlation of $+0.16$ between the ratio of height to weight and intelligence, showed a correlation of $+0.35$ between the morphologic index and intelligence. "No correlation was found in any one of the groups examined between height and intelligence. Similarly no correlation was found to exist between lung capacity and intelligence in 136 of the students making the group of 221, whose lung capacity had been measured. Weight gave a negative correlation."

Suggestive remarks are thrown out by the author to the effect that the morphologic type, (whether microsplanchnic, normosplanchnic or macrosplanchnic—it is the first that gives the more intelligent units) is the outcome of hereditary and accidental factors and that the microsplanchnic type corresponds to the hyperthyroid type; in this respect offering support to the common belief that the thyroid bears relationship to intelligence. Tables are given showing all measurements taken.

H. D. Kitson.

EDWARD K. STRONG, *The Psychology of Selling Life Insurance*.

Harper Bros., New York, 1922, pp 489.

The more immature a science is the more premature we are likely to be in attempting to apply it. The initial stage in almost every science has been marked by the belief that it could explain almost everything. Thus we have the climatic interpretation of civilization in the field of physical geography; the economic interpretation of history in economics; the Freudian interpretation of life in psychiatry, the 'glands regulating personality' in a new department of physiology; and even at this late date, the "time-binding" ability of man emanating from the staid realm of mathematics. So with the science of physics, there was a time when many people thought it could be used to explain all the phenomena of the universe. Today physics is one science among many, a great science to be sure, but nevertheless a science with distinct limitations. We do not find the newspapers and magazines full of 'The physics of this' and the 'physics of that.' We do not see among industrial literature frequent mention of the 'physics of production' or the 'physics of machinery.' And when a pitched baseball in its trajectory meets the head of a batter, people do not say: The ball hit him in the head at the *physical moment*.

What is not true about physics is distinctly true about psychology. We are quite likely to read on the sport page of a newspaper an article on "The Psychology of Training for a Prize-fight," or in any magazine an article on the "Psychology of Labor." Everything seems to have its psychological explanation and everybody seems to be "psychologically speaking." Why? Simply because psychology is in its early and undeveloped stage, which permits both psychologists and laymen to make claims and applications which will be considered absurd when it becomes a mature science. To be sure there will always be a common sense psychology, just as there is a common sense physics today. There have been common sense psychologists from the beginning of humanity. Such men as Christ, Socrates, Cicero, etc., understood human nature probably as well as we understand it today but they needed no psychological terms in which to describe their knowledge any more than the common sense physicist needs the scientific terms of physics to describe why he slipped on a banana peel or why his car skidded into another car. Similarly, the psychological terminology—and it consists mostly of the promiscuous use of the word psychology—in most cases adds nothing to the explanation of facts which could not be given just as well in ordinary common

sense language. The time will come when psychologists and laymen will no more think of saying 'the psychological moment' instead of 'the appropriate moment' than they do today of saying 'the physiological moment' when a man jumps up from a tack which he has accidentally sat upon.

The Psychology of Selling Life Insurance is a case in point. This book employs a certain number of psychological terms, but their use obscures rather than illuminates the lessons it seeks to inculcate. And unfortunately, it selects that part of psychology about which even psychologists admit there is the greatest doubt. The book is based, like so many others (e.g., *Instincts in Industry*, *The Psychology of Advertising*, *Influencing Men in Business*, etc.) on types of appeal, human motives—in short the instincts and impulses of human nature. Unquestionably, these forces exist, but psychology has added little that is new to our ability to use them. (We should except from this statement, the art of Psychiatry, which many psychologists seem to regard as an illegitimate offspring of psychology, and which does not enter into Dr. Strong's discussion.) Cicero in his day, could sway the Roman mob at will, though he had never seen a book on mob psychology. Today, when psychology has subjected human motives to the most minute analysis, there are probably not two psychologists in the world who agree on a classification of instincts, impulses or emotions. Some reduce all to two, some find as many as fifteen or twenty-five. As William James, whose word on this subject is probably more trusted than that of any other man, long ago said: "It is possible to make almost any classification of instincts on this basis, so long as it answers the writer's purpose." (*Principles of Psychology* II, 382).

When Dr. Strong begins to base his insurance talks upon such traits as 'submission,' 'approval,' 'gregariousness,' 'acquisitiveness,' 'love of wife and family,' etc., he is proceeding after a sound fashion, probably, but certainly he is contributing little from the field of psychology, simply because psychology in this field has next to nothing to contribute. If I were training salesmen in a knowledge of human motives, I should rather give them the Bible as a text for whatever our religious beliefs, the Bible presents the finest and most complete account of human motives and behavior ever written, both in respect to man's strength and his weakness. Whatever contribution Dr. Strong's book makes to the art of selling life insurance, it is certainly not in the field of psychology, even though its terminology may be drawn from text books in psychology. Judged from a common sense point of view, merely, it is packed with information and suggestions which ought to be exceedingly profitable to the insurance salesman.

and the style throughout is such as to make it one of the most readable business books which has appeared in a long time.

Henry C. Link
New York City.

WALTER LIPPMAN, *Public Opinion*.

New York, Harcourt, Brace & Co., 1922. pp. 427.

The thesis of this book is that the public, by which is meant the American citizenry, in order to govern themselves wisely, should be furnished with information about public questions.

The author laments the fact that such government research agencies as have operated, especially those that operated during the war, have not fulfilled their mission. The obstacles in the way of this are multitudinous requiring almost the entire 427 pages of the book for their recital. Some of them are physical, such as, lack of mediums of communication, and number of questions that effect public welfare. Others, chiefly of a psychological nature, are: the low intelligence of a great part of the public, which hinders them from distinguishing between the true and the false; the poverty of language; and the natural distortion involved in comprehending anything, which latter point the author develops in a popularly-phrased chapter on the psychology of perception. Finally there are social and economic obstacles consisting of vested interests, group and individual.

In the last four chapters the author describes the mechanism that he recommends for the amelioration of the ills under which public opinion and public action are now suffering. It is a series of Bureaus of Research and Information, one for each of the ten departments of government at Washington, with personnel independent of legislative and executive interference, and with funds insured against withdrawal. Such research organizations can be as well established with state and city governments.

The information gathered need not be disseminated throughout the entire nation. Few citizens would digest it if it were. It should, however, be placed at the disposal of expert special agents who can make use of it in preparing legislation or in executing the affairs of state.

With the admission of this as the main service of research bureaus, the author has shifted his ground. For whereas, during the first part of the book he plead for enlightenment of the whole citizenry, in the latter portion he apparently becomes frightened at the stupendousness of the undertaking and uses his panacea for a different end—the enlightenment of a selected few.

This latter aim, as every one must admit, is quite feasible;

though merely to enlighten public officials does not insure disinterested and statesmanlike conduct on their part. Still, to give them a fount of scientifically gathered information will probably increase the chances of high-grade management of public affairs. And though not capable of influencing public opinion directly, bureaus of research may do so indirectly by bringing to light bases of intelligent judgment. And though to effect intelligent direction of public affairs very soon were too much to expect in view of the proneness of humans to act upon feeling, still the goal is a righteous one; and the collection of information is probably the point at which to begin.

H. D. Kitson.

A. M. SIMONS, *Personnel Relations in Industry*.

The Ronald Press, N. Y. 1921, pp 341.

This book enumerates the problems of personnel administration usually treated by books upon the subject, and discusses them with adequate expertness and clearness.

The psychologist-reader will be gratified, on perusing the thirty or forty pages devoted to the contributions which psychology can make to the field, to find that the author has sought the correct sources for most of his psychologizing, and has tapped them discriminatingly. The pages are heavily annotated, as the author seeks confirmatory authority for his statements by numerous quotations, thus providing a good bibliography.

The pages also smack of the class-room, by reason of a noticeable outline method of presentation—the book grew out of the author's Extension Lecturership in the University of Wisconsin—but this, while detracting from the smoothness of style, will be regarded as a pedagogical asset by those who use the book as a class text. It is, however, not objectionably academic. It bears the imprint of contact with actual business practice.

The books upon personnel relations and allied subjects are so numerous and widely distributed among personnel executives that experienced personnel workers will hardly find anything new in this presentation. Still the large number of business executives and intelligent laymen who have not studied personnel problems will probably find here a good deal of information that is new to them, information that is thoughtfully considered and interestingly presented. Even well-enlightened executives may find in the chapter on Wider Social Control the names of a number of agencies—public and private—for fostering desirable personnel relations that they have never heard of.

H. D. Kitson.

NEW BOOKS AND PAMPHLETS RECEIVED¹

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio, or W. F. Book, Indiana University, Bloomington, Indiana.

- GAMBRILL, BESSIE E. *College Achievement and Vocational Efficiency*. Columbia University Contributions to Education, No. 121. Teacher's College, New York City, 1922.
- KNIGHT, FREDERICK BUTTERFIELD. *Qualities Related to Success in Teaching*. Columbia University Contributions to Education, No. 120. Teacher's College, New York City, 1922.
- RUSSELL, CHARLES. *The Improvement of the City Elementary School Teacher in Service*. Columbia University Contributions to Education, No. 128. Teacher's College, New York City, 1922.
- MACDOUGALL, ROBERT. *The General Problems of Psychology*. The New York University Press, New York City, 1922.
- MICKLEM, F. R. *Miracles and the New Psychology*. A study in the Healing Miracles of the New Testament. Oxford University Press, American Branch, 1922. pp 142. Price \$2.50.
- TEN BROEKE, JAMES. *The Moral Life and Religion*. The Macmillan Co., New York. pp. 244.
- ROGERS, ARTHUR KENYON. *The Theory of Ethics*. The Macmillan Co., New York, 1922. pp. 197.
- JONES, RUFUS M. *Spiritual Energies in Daily Life*. The Macmillan Co., New York, 1922. pp. 179.
- MCCONNEL, REV. S. E. *Confessions of an Old Priest*. The Macmillan Co., New York, 1922. pp. 125.
- BINGHAM, HAROLD C. *Visual Perception of the Chick*. Behavior Monographs, Vol. IV, No. 4, 1922. Williams & Wilkins Company, Baltimore, Md. 104 pp.
- BOBBIT, FRANKLIN. *Curriculum-Making in Los Angeles*. Supplementary Educational Monographs. No. 20. The University of Chicago, Chicago, Ill. 106 pp.
- BUSWELL, GUY THOMAS. *Fundamental Reading Habits: A Study of Their Development*. Supplementary Educational Monographs. No. 21. The University of Chicago, Chicago, Ill. 150 pp.
- CLEVELAND FOUNDATION. *The Professional Education of Teachers in Cleveland*. The Board of Education, Cleveland, Ohio, June 1922. 93 pp.
- CLEVELAND PUBLIC SCHOOLS. *Junior High School Mathematics*. Board of Education, Cleveland, Ohio, 1922. 45 pp.
- CLEVELAND PUBLIC SCHOOLS. *Surveys of the Department of Instruction, Cleveland Public Schools*. Board of Education, Cleveland, Ohio. 29 pp.

¹Mention here does not preclude further comment.

- COBURN, CHARLES A. *Heredity of Wildness and Savageness in Mice*. Behavior Monographs, Vol. IV, No. 5, 1922. Williams & Wilkins Company, Baltimore, Md. 71 pp.
- Consolidation of Schools in Maine and Connecticut*. Department of the Interior, Rural School Leaflet No. 4, May, 1922. Washington, D. C. 12 pp.
- DEFFENBAUGH, W. S. *Administration of Schools in the Smaller Cities*. Department of the Interior, Bulletin No. 2, 1922. Washington, D. C. 75 pp.
- DOWNEY, JUNE E. *Downey Group Will-Temperament Test*. World Book Company: Publishers, Yonkers-on-Hudson, New York. 12 pp.
- Educational Research Bulletin*. Pasadena City Schools. Vol. I, No. 1. Department of Educational Research, Chamber of Commerce Building, Pasadena, California. 6 pp.
- GINSBERG, MORRIS. *The Psychology of Society*. E. P. Dutton & Company, 681 Fifth Avenue, New York. 174 pp.
- GRAY, WILLIAM SCOTT. *Remedial Cases in Reading: Their Diagnosis and Treatment*. Supplementary Educational Monographs, No. 22. The University of Chicago, Chicago, Ill. 208 pp.
- Home Economics in Rural Schools*, Department of the Interior, Home Economics Circular No. 13, May 1922. Washington, D. C. 15 pp.
- INGLIS, ALEXANDER. *Intelligence Quotient Values*. World Book Company, Yonkers-on-Hudson, New York.
- LOMBARD, ELLEN C. *Home Education by Means of Reading Courses and the Cooperation of State and National Agencies*. Department of the Interior, Home Educational Circular No. 2, April 1922. Washington, D. C. 11 pp.
- KING, LE ROY ALBERT. *Status of the Rural Teacher in Pennsylvania*. Department of the Interior, Bulletin No. 34, 1921. Washington, D. C. 87 pp.
- List of References on Rural Life and Culture*. Department of the Interior, Bureau of Education, Library Leaflet No. 16, May, 1922. Washington, D. C. 10 pp.
- LOOSMORE, W. CHARLES. *The Gain of Personality*. E. P. Dutton & Company, 681 Fifth Avenue, New York. 233 pp.
- MEANS, MARIE HACKL. *A Tentative Standardization of a Hard Opposites Test*. Psychological Monographs. Psychological Review Company, Princeton, N. J. 64 pp.
- MUNSTERBERG, MARGARET. *Hugo Munsterberg: His Life and Work*. D. Appleton and Company, New York. 448 pp.
- PRESSEY, SIDNEY L. AND LUELLA COLE. *Introduction to the Use of Standard Tests*. World Book Company, Yonkers-on-Hudson, N. Y. 263 pp.

- Reorganization of Home Economics in Secondary Schools.* A report of the Commission appointed by the National Education Association. Department of the Interior, Bulletin No. 5, 1922. Washington, D. C. 38 pp.
- Report on the Higher Educational Institutions of Arkansas.* Department of the Interior, Bulletin No. 7, 1922. Washington, D. C. 18 pp.
- Review of the Educational Legislation, 1919 and 1920.* Department of the Interior, Bulletin No. 13, 1922. Washington, D. C. 30 pp.
- Statistics of Teachers' Colleges and Normal Schools, 1919-20.* Department of the Interior, Bulletin No. 8, 1922. Washington, D. C. 68 pp.
- Status of Sex Education in High Schools.* Department of Interior Bulletin No. 14, 1922. Washington, D. C. 12 pp.
- TERMAN, DICKSON, SUTHERLAND, FRANZEN, TUPPER, FERNALD. *Intelligence Tests and School Reorganization.* World Book Company, Yonkers-on-Hudson, New York. 111 pp.
- TERRY, PAUL WASHINGTON. *How Numerals Are Read.* Supplementary Educational Monographs, No. 18, June 1922. The University of Chicago, Chicago, Ill. 109 pp.
- WALLIN, J. E. WALLACE. *The Achievement of Subnormal Children in Standardized Educational Tests.* Teachers College, Miami University, Oxford, Ohio. (Miami University Bulletin, Series XX, No. 7, April 1922.) 97 pp.
- WELLS, GEORGE ROSS. *Youth and The Open Door.* E. P. Dutton & Company, 681 Fifth Avenue, New York. 175 pp.
Mention here does not preclude further comment.
- BOOK, W. F., A. S. BEARD, and O. R. CHAMBERS. *Studies in Observational Learning.* Preliminary reports on Group Tests for measuring observational learning and accuracy of report; Educability of Observation and Report; Relation of Intelligence and School Training to Observational Learning and accuracy of Report. Bulletin of Extension Division, Indiana University, Vol. VII, No. 12, Aug., 1922. pp. 90.

THE JOURNAL OF APPLIED PSYCHOLOGY

VOL. VI

DECEMBER, 1922

No. 4

MEASURES OF GENERAL INTELLIGENCE AS INDICES OF SUCCESS IN TRADE LEARNING.

BY KARL M. COWDERY

Whittier State School

Psychological tests are being widely investigated as a means of identifying persons who have the abilities to make successful records in industrial and mechanical occupations. For this purpose several types of tests have been suggested, among which are analytical tests such as the series used by Dr. Helen T. Wooley (18)¹, which seek to measure specific abilities; the "sample" tests of the performance trade test type such as are described by J. Crosby Chapman (4), wherein the subject performs operations which actually appear in the field of activity for which the person is applying; the analagous tests, described by H. E. Burt (3), in which the person tested is made to perform operations which are considered as containing the elements involved in the occupation; and the miniature which simulates on a small scale the exact conditions of the commercial task. In his book entitled "Employment Psychology," H. C. Link (10) describes the methods and applications of psychology to employment in both general and specific industries, and outlines the extent to which psychological methods are and may be used. He implies in his chapter on General Intelligence that he finds little or no relation between general intelligence and industrial abilities but states that specific abilities are of more value as indices of the potential capacity of a prospective employee. He is dealing, however, with operations rather than with occupations or trades.

Other writers who have considered the problem of ability for trade learning, in its broadest sense, have not agreed among themselves as to the relations which exist between measures of general

¹A study made under the direction of Dr. Lewis M. Terman, Stanford University, and Dr. J. Harold Williams, Director, California Bureau of Juvenile Research, Whittier State School, Whittier, California.

²Numbers in parenthesis indicate references cited at the end of this study.

intelligence and success in occupations. The lists of occupations open to the feeble-minded given by Miss M. L. Anderson (1) and by Miss Ada Fitts (5), and the industrial capacities outlined by Dr. H. H. Goddard (7), indicate that persons even of low intellectual standing are not entirely barred from useful vocational activity, particularly if given suitable supervision. Dr. George Ordahl (11) gives figures which, within a certain small range and at a low level of mental rating, suggest that morons can engage in productive labor along with those of higher abilities provided their work is properly planned and supervised. He also suggests that for this work (tomato canning) the persons of higher levels of intelligence are proportionately more productive. C. S. Rossy (12) states that his study of institutional work suggests farm work most suited to the feeble-minded. Miss Anderson (1) points out that the work which is available for the low grade individuals is not trade work and that defectives cannot take part in the more skilled trades demanding shop training. The question then arises, What is the lowest limit of mental rating that will permit of trade training?

Assuming that the lowest limit of intelligence permitting successful trade learning has not been found, what is the relation of success in learning to the extent to which the learner's rating exceeds the minimum level? Bridges and Coler (2) find that the children of the skilled class of workmen have average intelligence levels distinctly higher than those of the unskilled laborers. Greater progress, in these cases, has been made by the parents of children whom tests have shown to possess superior intelligence. If the relative levels of intelligence among the children fairly represent the status of the parents the conclusion is to be drawn that greater general mental ability foretells more advanced occupational learning. Dr. L. M. Terman (13) and Dr. Helen T. Wooley (17) find similar differences between different classes of workmen. These studies relate to associated problems but do not measure directly the relation of general intelligence to learning the same tasks or to success in the same activities.

J. K. Flanders (6) found a low correlation between I. Q. and the salaries paid to express company employees and concluded that there are other factors of more importance than intelligence which determine the value of a man to his employer. M. S. Viteles (16) expresses the opinion that "trade competency is related to specific abilities which the individual possesses," and that "general intelligence tests can not be used for the selections for a great mass of skilled and semi-skilled jobs, such as office and clerical jobs." On the other hand Dr. L. L. Thurstone (14) says "I do not believe that

office work has any special abilities," and as a test combines several psychological measures, finding a combination whose scores correlate more highly with the rated abilities of the subjects than do any of the individual tests for specific abilities. R. W. Kelly (9) believes that intelligence tests are most useful in connection with upper levels of work such as clerical workers, executives and salesmen. Link (10) found a correlation of .55 between general intelligence and technique in office work. For ability in specific tasks such as the gauging and inspection of shells he found but low correlations between mental rating and ability to do that work. The U. S. Army (15) used mental rating as one of the important criteria in selecting officer training material, and also found, according to Henmon (8), direct relations between native intelligence and flying ability in the aviation training fields.

Consideration of the material in the literature cited suggests two problems in connection with the task of selecting trade training for boys, (a) the minimum necessary level of intelligence for successful learning of the activities and coordinations involved in the various trades, and (b) the relative value of different levels of intelligence above the minimum.

At the Whittier State School, Whittier, California, boys between the ages of 12 and 19 have been assigned to the learning and performing of "trade" work under conditions which may be termed apprenticeship. That is, they are working with the materials and in the operations of trades directly under the guidance and supervision of experienced journeymen or expert workmen in the respective trades. The assignments to the different trades have been such that a considerable range of mental age level is represented in every group of boys in the various "details."

This study was undertaken to determine with more exactness than is afforded by routine observation to what extent the measure of the general intelligence of the boys can or should be used in deciding the specific assignment of a boy to one trade rather than to another. The objectives or criteria sought in the study are those suggested above, namely, the minimum level of intelligence necessary for success, and the relation of success to varying degrees of mental ability for the different trade groups.

For such a study there were available to the writer the information contained in the files of the Whittier State School and those at the California Bureau of Juvenile Research¹ which gave the results of individual tests of each boy by means of the Stanford Re-

¹The writer is indebted to Superintendent Fred C. Nelles of the Whittier School and to Dr. J. Harold Williams, Director of the Bureau of Juvenile Research at the Whittier State School, for their co-operation in making available the material for the study on which this article is based.

vision of the Binet-Simon Intelligence tests. The data used include the I. Q., the age of the boy at the time of assignment to trade work, and the mental age at that time.

As a measure of the success of the individual boys in their respective trade work only a descriptive report of the boys' work was available at the time of starting the study (1919). A new form of report was then prepared for use thereafter whereby at the beginning of each month the instructor made a rating of the work done by each boy during the preceding month. Three years of accumulation of these records have given a set of ratings on each of 578 assignments to 22 different trade groups.

The ratings of the instructors were on a five division scale whereby the work of the boy was scored as *excellent*, *good*, *fair*, *poor* or *bad*. To obtain a single index of the work done by the boy during his entire period in the trade an arithmetical average of all ratings for each boy was made, known in this study as the Trade Average.

To find the relations which actually existed between the success of the individuals and their mental ratings, distributions for each group were made of the Trade Averages by Mental Age Levels. From these distributions the coefficients of correlation were computed and graphic representations were made to determine the critical levels of minimum ability required for success. As suggested by H. O. Rugg (Statistical Methods

TABLE I
DISTRIBUTION OF INTELLIGENCE QUOTIENTS

I. Q. s	f	%
1.25-1.29	2	0.3
1.20-1.24
1.15-1.19	5	0.8
1.10-1.14	3	0.5
1.05-1.09	22	3.8
1.00-1.04	26	4.5
.95-.99	49	8.5
.90-.94	61	10.6
.85-.89	82	14.2
.80-.84	80	13.9
.75-.79	83	14.4
.70-.74	65	11.2
.65-.69	53	9.2
.60-.64	21	3.6
.55-.59	16	2.8
.50-.54	6	1.0
.45-.49	4	0.7
Totals	578	100.0

* Indicates level of median case.

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 315

Applied to Education, p. 291) correlations were made by the Spearman Footrule Formula, or method of grades, where the number of cases is less than 30. For frequencies of more than 30 the Pearson Products Moment Formula was used. The critical levels were so located that 75 per cent. or more cases below that level fail to receive a rating of median or above, or that 75 per cent. or more of the cases above that level are rated as median or better. The former represents the "lower critical level" or minimum level necessary for success. The latter is an indication of a level above which there is greatest expectation of success.

Correlations of success with I. Q. and with Chronological age are computed for comparison. Table I gives the distribution of the

TABLE II

DISTRIBUTION OF CHRONOLOGICAL AND MENTAL AGES.

Levels	f C.A. %	f M.A. %
19:6-19:11	3 0.5	
19:0-19:5	2 0.3	
18:6-18:11	1 0.1	
18:0-18:5	3 0.5	
17:6-17:11	8 1.4	2 0.3
17:0-17:5	17 3.0	3 0.5
16:6-16:11	39 6.8	5 0.8
16:0-16:5	72 12.5	5 0.8
15:6-15:11	91 15.8	13 2.3
15:0-15:5	95 * 16.5	17 3.0
14:6-14:11	88 15.3	31 5.4
14:0-14:5	68 11.8	39 6.8
13:6-13:11	52 9.0	48 8.3
13:0-13:5	21 3.6	53 9.2
12:6-12:11	9 1.4	49 8.5
12:0-12:5	4 0.7	67 11.7
11:6-11:11	3 0.5	46 * 8.0
11:0-11:5	2 0.3	40 6.9
10:6-10:11		53 9.2
10:0-10:5		37 6.4
9:6- 9:11		24 4.2
9:0- 9:5		14 2.4
8:6- 8:11		18 3.1
8:0- 8:5		7 1.2
7:6- 7:11		5 0.8
7:0- 7:5		1 0.1
6:6- 6:11		1 0.1

Totals 578 100.0 578 100.0

* Indicates level of median.

Intelligence Quotients for the 578 cases. Table II is a tabulation of the chronological and mental ages of the boys at the time each entered the trade group. The medians of the actual and mental

ages are noted as being 15:3 and 11:9 respectively. For the group there is an average retardation of 3 years and 6 months, the median I. Q. being .82.

Table III gives the frequencies for each trade group together with the coefficients of correlation between the Trade Averages and the Intelligence Quotients, chronological ages and mental ages.

TABLE III
CORRELATIONS WITH TRADE AVERAGES.

Trade Groups	f	I. Q.	C. A.	M. A.
I. Printing Trades				
Print Shop	51	.116±.092	.314±.086	.334±.084
Bookbindery	18	-.200±.145	-.005±.15	-.305±.14
II. Mechanical				
Carpenter	32	.255±.112	-.088±.119	.306±.109
Paint	20	.283±.138	-.366±.13	.283±.138
Plumbing	11	-.259±.14	-.045±.15	-.377±.13
Blacksmith	21	.211±.141	.244±.138	.299±.134
Garage	11	.259±.16	.259±.16	.259±.16
III. Garment				
Tailor	78	-.212±.073	.227±.072	.232±.072
Shoe	34	-.342±.102	-.003±.117	-.313±.103
Laundry	33	-.308±.106	.015±.117	-.297±.107
IV. Culinary				
Kitchen	32	-.343±.105	.600±.067	.093±.120
Bakery	26	-.275±.122	.476±.102	-.016±.132
V. Farming				
Dairy	32	-.147±.117	.091±.119	-.126±.118
Poultry	22	.337±.127	.393±.122	.600±.092
Teamsters	24	-.013±.138	-.105±.136	-.087±.137
Vegetable G.	20	-.184±.146	.307±.137	-.234±.143
VI. Miscellaneous				
Office Boys	7			.983±.009
Flower Garden	18	-.169±.146	-.169±.146	-.071±.150
Power House	8	.000	.000	.247±.224
Hospital	7		.414±.211	.414±.211
VII. Service				
Housekeepers	38	-.094±.108	-.146±.107	-.147±.107
Dining Room	35	.055±.114	.001±.114	.030±.114

Charts I, II, and III are sample diagrams representing the three types of relationship found between Trade Averages and mental levels. Each of these diagrams is discussed in connection with the trade group which it represents.

PRINTING TRADES

1. **Print Shop.** The work of the print shop includes the production of record forms, stationary, and other similar material for the use of the School offices; a bi-weekly School paper of local news and interests; a bi-monthly scientific journal, the *Journal of Delinquency*; and various bulletins, monographs and other less ex-

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 317

tensive pamphlets both for the school and for the Bureau of Research.

Training and experience are provided in the theory and practice of type-setting (hand method), paging, make-up, press-feeding, and the various processes involved in job printing. The average size of the class is 14 boys.

The instructor has had commercial experience in printing as a "trade" without the limited specialization common to the activities of the occupation. In addition to his commercial work he has taken teacher-training class instruction toward certification as a vocational teacher.

The distribution of ratings, given in Chart I, has made use of the full rating scale and has closely approximated the normal distribution curve.

Correlation represented by a coefficient $.334 \pm .084$ between mental age levels and Trade Averages is found. This is interpreted as meaning that there is a definite, tho not large, tendency for probability of success to be relatively greater as the standing of the individuals receiving instruction approaches and exceeds the median intelligence of the group. The diagram of the distribution shows that on the basis of 75 per cent or more success an upper critical level is justified at the mental age of 15:0. In other words, of the 12 cases rated as having mental ages above 15 years, 9 receive median or better Trade Averages. Similarly at the other end of the range, 6 of the 7 cases below a mental age of 12:3 fail to reach the median of success. The lower critical level is placed at this location.

The conclusion is drawn that a minimum mental age level of 12 years 3 months is requisite for reasonable progress in learning the printing trade; that more and more success is expected as higher mental levels are represented, and that of those cases above the level 15:0 at least 75 per cent are expected to be successes.

The coefficient of correlation, $.314 \pm .086$, between the Trade Averages and chronological age levels, suggests that the factor of actual age should also be considered in making selections of boys for this trade.

A partial correlation was computed to find the relation of mental age to trade success, eliminating the factor of chronological age and the change was so slight ($-.06$), being less than the Probable Error of the correlation, that this element of the distribution of actual ages at the various mental ages was considered as negligible.

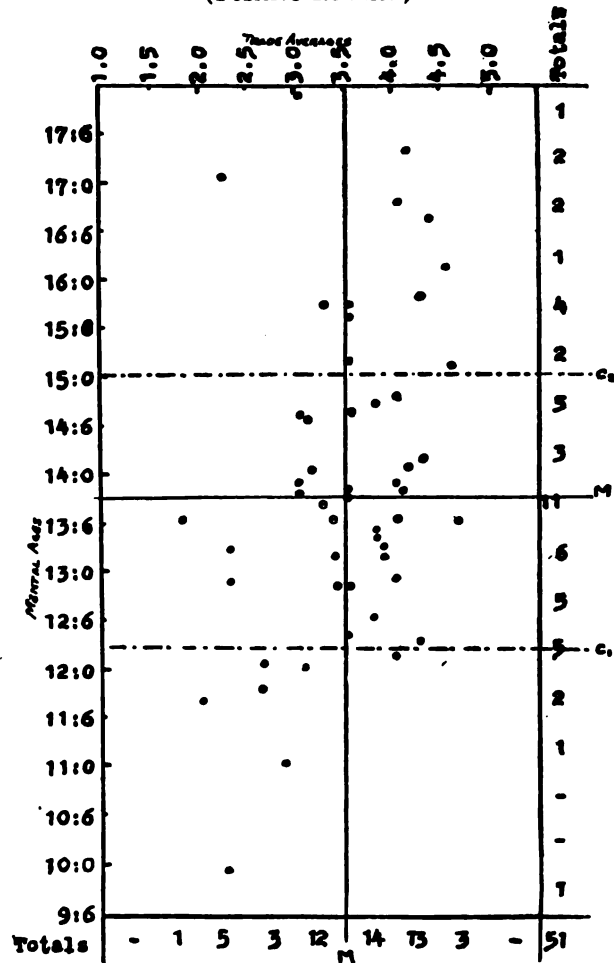
2. Book-binding. In the associated trade of book-binding the products of the print shop are folded, trimmed, bound or padded as

COWDERY

CHART I

PRINT SHOP

Distribution of Trade Averages by Mental Age Levels
(Positive Relation)



LEGEND
M Median
C_u Upper Critical Level
C_l Lower Critical Level

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 319

may be required. Old books are rebound, and small leather articles such as purses and notebook covers are manufactured. There is a comparatively small range of activities involved, and the group of boys is small, four being the average size of the group. This allows for close personal supervision of the work. The combination provides for repeated processes under supervision closely resembling factory conditions.

The relation between mental level and Trade Averages is found to be negative, coefficient of correlation $-.305 \pm .14$. The boys of lower mental ages adapt themselves best to the type of work involved and the learning of the process requires no more than limited intelligence. The meaning of the lower critical level in this distribution is reversed from that in the print shop group. Seven of the 8 boys having mental ages below 12 years receive averages of median rank or better.

The relation of actual age to success is practically neutral, hence the suggestion is that boys of inferior intelligence are suitable for this kind of work within the limits of the actual ages represented, 11:2 to 15:8.

The range of mental ages for this group did not go below 10 years. It is quite likely that at some level lower than this a critical level would be found which would represent the minimum intelligence level required. We are safe in saying that it is less than 10 years of mental age.

MECHANICAL TRADES

1. Carpenter Shop. The Carpenter Shop gives supervised experience in work which includes the manufacture of furniture and other articles demanding training in "cabinet-making," as well as experience in rougher types of carpentry, and the operation of mill machinery. The shop takes care of all repair work for the school, both of furniture and buildings.

The instruction is in charge of a tradesman with experience in both cabinet-making and mill work, who also holds a special certificate for vocational instruction.

The relation between mental age levels and trade ability is expressed by the co-efficient $.306 \pm .109$. The association is positive, indicating that intelligence is a factor but not the sole factor which determines degree of success in learning the activities of the trade.

In spite of the positive relation indicated by the coefficient the range of mental ages failed to reveal any critical levels. Within the range 10:0 to 15:0, there is a slight tendency for more success at the upper end of the range than at the lower. Also it is quite no-

ticeable that of those with mental ages below 12:0, the deviations from median success of the failures are greater than those of the successes. In other words, with about even numerical proportions of successes and failures, the degree of the success of those below the median mental age is comparatively limited.

Although boys of mental ages as low as 10 years can learn, those individuals whose intelligence rates above 12:3 attain to greater degrees of success.

2. **Paint Shop.** The work of the painting detail consists of the finishing of furniture, including flat painting, staining, varnishing, and rubbed finishing; interior painting and decorating, and floor surfacing; outside painting; auto body work; curtain-hanging and glazing. The boys are given instruction in this work, as well as in the preparation of materials, color mixing, care of brushes and tools, and some training in sign construction and lettering.

The instructor is a painter of considerable trade experience and holds a special certificate for vocational instruction.

For the group of 20 boys the distribution presents a positive relationship between mental and trade ability, correlation coefficient $.283 \pm .138$. The tendency is small but definite. The diagram suggested both an upper and a lower critical level at 14:0 and 11:0 respectively. Not a single median rating was attained by any boy whose mental level was below 11 years. Four out of five above the upper level were rated at the median or better.

The negative relationship between chronological age and success suggests that boys can most profitably be started in the learning of this trade before they are 16 years of age.

The minimum mental age of 11 years is required for successful learning of the operations involved and the highest proportion of success will be expected of boys between 14 and 16 years old whose mental levels are 11 years or higher, preferably above 14 years.

3. **Plumbing Shop.** The plumbing shop duties consist of general repair work on water, gas, and steam service lines, with little new installation work. The instructor is a journeyman plumber in charge of from one to three boys at a time. He has had no special instructional training.

The limited number of boys observed renders unreliable any statistical indices. The eleven cases rated gave a negative coefficient of correlation between mental age levels and Trade Averages, $-.377 \pm .13$. No general conclusions seem to be warranted although this would seem to place this group in the class of trade where intelligence is not an important factor in the early learning process.

4. **Blacksmith and Machine Shop.** This detail handles all of the

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 321

metal working for the school, including forge work and a limited amount of machine practice in the use of lathes, shapers, milling machines and drill presses. The instructor is a forge man who has had some additional machine operation.

The coefficient $.299 \pm .134$ represents a limited positive relation between intelligence and success. A lower critical level is suggested at M. A. 12:6, but the cases below that level are barely inferior to those attaining median success. Without more cases this level is questionable. However the highest degree of success is attained by those whose mental levels are above 13:6.

Similarly a chronological age of 15:9 or higher predicts the most likelihood of efficient learning.

5. Garage. The work of the boys in the garage detail consists primarily in the observation of repair work and actual practice in the general maintenance and driving of automobiles and motor trucks. The detail is small and renders questionable the meaning of a positive correlation coefficient of $.259 \pm .16$.

GARMENT TRADES

1. Tailor Shop. The operations performed in the tailor shop come under three heads, (1) the mending and darning of worn and torn garments, (2) the sorting and inspection of laundered clothing and its issuance to the boys of the School, and (3) the making of new underwear, shirts, work trousers, caps and suits. The work is under the supervision of a woman whose experience in the trade has been largely that of institutional tailoring instruction. The detail contains an average of 20 boys.

The ratings of the instructor are contained in a comparatively narrow range of the rating scale and are therefore not as reliable as the distribution represented in the charts. ✓

For the 78 cases included in this detail a slight positive relation is found, $.232 \pm .072$. But in the wide range of mental ages, 7:6 to 17:8, no critical levels are indicated. Apparently very low intelligences are sufficient for productive work in this shop, and the lower critical level falls below 7 years 6 months.

2. Shoe Shop. Shoes are both made and repaired in the shoe shop. A few other leather articles such as footballs are cut and made from pattern. Each boy passes in succession thru the various processes involved in the making of the shoes, but since the styles of shoe are limited, the number and variety of problems faced are comparatively few. There is considerable repetition of activity.

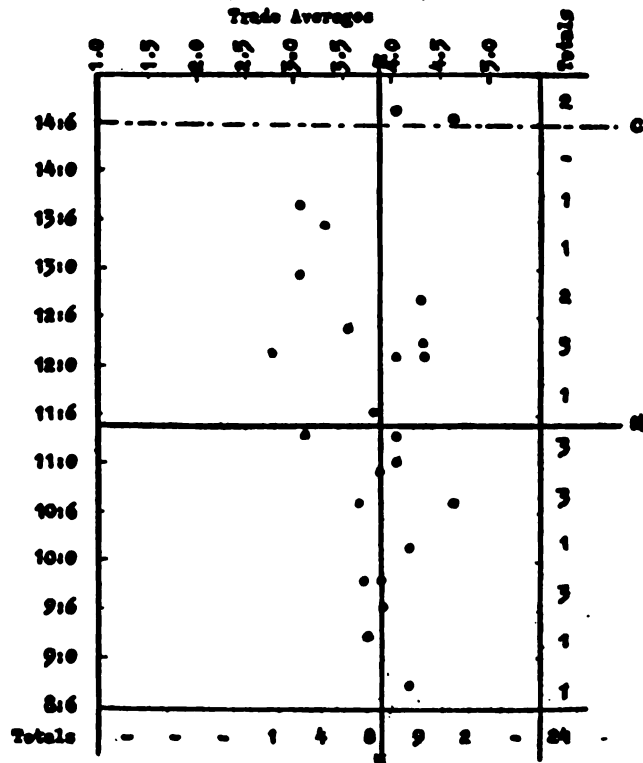
A distinct negative relationship is found between mental rating and Trade Averages, $-.313 \pm .103$. Only a very low level of intelligence is required for successful learning, the minimum level

falls somewhere below the lower limit of our range, 6:6. On the other hand the higher levels of ability find difficulty in adapting themselves to the operations and restricted supervised activities of the trade learning under the shop conditions. Chart III represents the distribution and shows the reversed critical levels at years 10:0 and 12:3.

CHART II

TEAMSTERS

Distribution of Trade Averages by Mental Age Levels
(Neutral Relation)



LEGEND

M Median

C Suggested Possible Critical Level

3. Laundry. The work of the boys in the laundry consists of hand ironing and attendance upon the comparatively simple machines in use in modern laundries. The instructor supervises much

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 323

as does the foreman in commercial work, and also cares for the up-keep and operation of the more complicated machines. Under these circumstances the group assigned to the detail has been of lower average intelligence than those working in the previously discussed groups, median mental age 11:10. The range, however, has been fairly comprehensive, 8:0 to 16:5. For the group the correlation of mental age with trade success is $-.297 \pm .107$, and inspection of the distribution revealed consistent failure for the boys whose mental ages were above 15 years. Likewise the successes are most frequent among those below the 10:6 level. The critical levels are reversed and definite. A minimum requirement of intelligence is not defined within the range represented.

CULINARY TRADES

1. Kitchen. All food except the bread and pastries made in the bake shop is prepared for serving by the kitchen. For purposes of segregation and because of the commercial opportunities after leaving the school, this detail is made up largely of negro boys.

A range of intelligence levels from 7:6 to 15:8 is included in the group, with a median mental age of 12:3. The chef failed to make use of the larger part of the rating scale, with the result that the lowest Trade Average is above the median of the ratings given by the other instructors. It is questionable whether every boy in the detail made the success indicated. However variations within the group are present altho in very slight amounts. The resulting coefficient of correlation, $.093 \pm .12$, is practically zero. There seems to be little if any relation between intelligence and comparative success. However, at each extreme of the range there is a group of which the majority fail to attain to the median rating of the group. Above mental age 13:3, not a single case reached the median rating, while a similar distribution occurs below the level 9:6. Greatest success occurs between the mental ages 9:6 and 13:3, but the lowest rating as noted above is comparatively high.

2. Bake Shop. In the bake shop are prepared the bread and pastries (pies, cakes, cookies, doughnuts, etc.) for the dining rooms. The boys do the work of mixing the doughs, timing the rising and baking processes, and remove the completed articles from the ovens; all under the supervision and instruction of a first-class baker.

Here again the relation between mental rating and Trade Averages is practically zero, $.016 \pm .132$, altho the range of rating is complete. The factor of intelligence is nearly negligible under the conditions of supervision provided in this shop.

On the other hand chronological age is a factor to be used as a

guide in the selection of boys, the coefficient of correlation with Trade Averages being $.476 \pm .102$. A chronological age of 14:6 is required for reasonable expectation of success and both higher proportions and degrees of success are attained by the still older boys.

FARM WORK

1-2. **Dairy and Teamsters.** The similarity of the statistical indications and of the diagrammatic distributions for these two details makes possible discussion of the two groups together after the duties of each have been mentioned.

The work in the dairy consists in the care of the physical equipment, barns, stalls, milk receptacles, feed and feed-bins, of the feeding of the cows, and handling of the milk after cooling and separation. The milking is done by hired milkers. Assignments are made from the dairy detail to the feeding of the hogs and the care of their pens.

The teamsters are supervised in the care of horses, wagons, and the horse barns. They drive the teams for all hauling service not handled by motor trucks, for plowing, cultivating, etc., where tractor operation is not feasible. All tractor work is done by hired help. In two cases the tractor operators have been boys who have formerly been members of the teaming detail and who have been hired after special instruction in tractor operation.

In the case of each of these groups the relation of mental age to subsequent success in the work is very small and negative (practically zero, in view of the P. E.s). Coefficients are for the dairy $-.126 \pm .118$, and for the teamsters $-.087 \pm .137$. The distribution of Trade Averages by mental ages consist of nearly even divisions of successes and failures at all levels from mental age 8:0 to 14:6. Chart II gives the distribution of teamster ratings and illustrates the mental relationships. In each group all boys with mental ages above 14:6 receive better than median averages. This suggests a possible upper critical Level at this point. Likewise the relation between chronological age and success is zero.

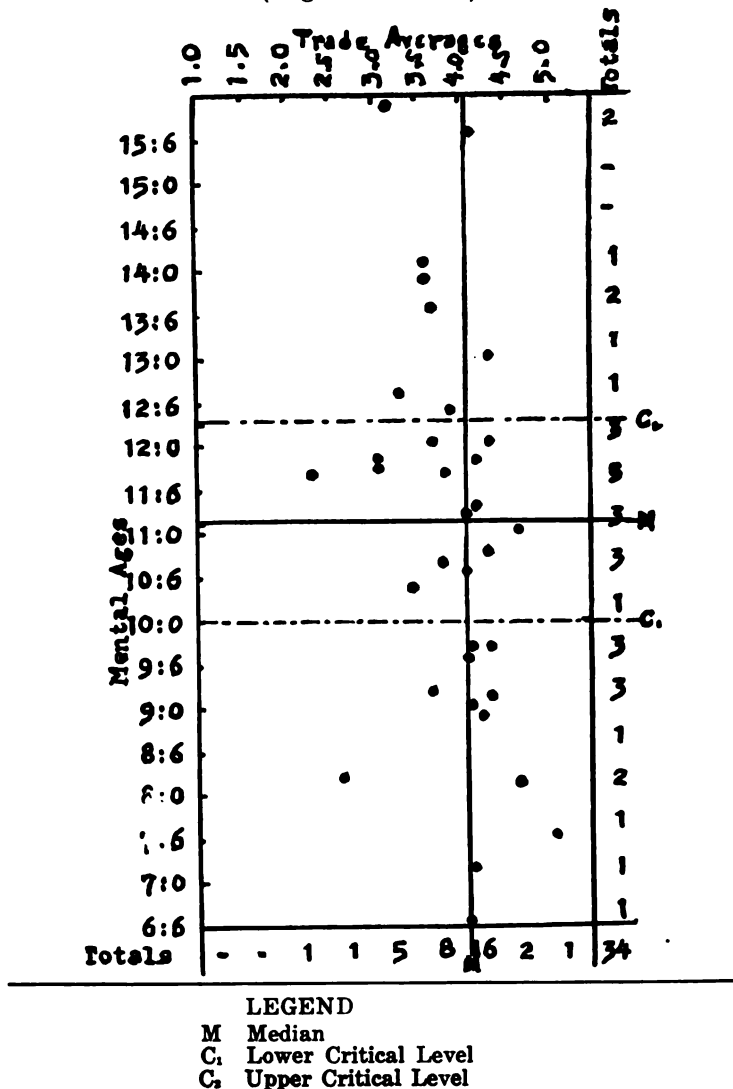
The practical suggestion is that boys between ages 13 and 19 and whose mental ages are above 14:6 can be expected to make successful progress in this work; and that for mental ages below 14:6 the expectation is about equally divided between success and failure. The minimum intelligence required is below the level of 8 years.

3. **Poultry Department.** Instruction is given in the feeding of poultry, breeding of pure strain stock, care of pens, handling of egg-laying and records, incubation and hatching. The instructor is an experienced chicken fancier and egg producer, but has had no training in educational methods.

CHART III

SHOE SHOP

Distribution of Trade Averages by Mental Age Levels
(Negative Relation)



The relation found between the ratings and the mental levels of those in the group shows the highest positive correlation, except one, of all the groups studied. The coefficient is $.60 \pm .092$.

Two clear cut critical levels are indicated at the 12:6 level and the 10:6 division.

There is a similar but not so marked relationship between actual age and success.

The most effective learning is achieved by boys of the higher ages whose mental levels approximate or exceed 12 years 6 months.

4. Vegetable Gardens. Supervised care of plant propagation and cultivation is the field of activity of this detail. All planning of planting is done by the instructor and farm manager. The boys are essentially supervised workers.

For this group a negative correlation, coefficient $-.243 \pm .143$, is found. Most satisfactory results are obtained from the boys of the lower levels of intelligence. As suggested by Rossy (12), this type of work is especially adapted to defectives.

The positive correlation .307, altho small, suggests chronological age rather than mental ability as a criterion for selecting workers in this detail.

MISCELLANEOUS UNCLASSIFIED

1. Office Boys. Altho a very small group of boys has been assigned to this work of running errands for the administrative offices of the School, the distribution of their ratings is interesting as indicating almost perfect correlation with mental level. Correlation coefficient $.983 \pm .009$. The boys who made median or better success were all of mental ages above 13 years 3 months.

2. Flower Garden. The boys of this group are assigned to routine tasks such as hoeing weeds, trimming bushes, picking flowers, but under comparatively lax supervision. The emphasis in the selection of boys placed on dependability rather than upon other factors seems to be at least partially warranted by the relation found between two other factors and trade ratings. The correlation with mental age, $-.071 \pm .150$, and that with chronological age, $-.169 \pm .146$, are both practically zero and seem to be almost negligible in selection considerations. The evidence is insufficient to validate a critical level suggested at mental age 14:6. Both of the two cases above that level attain to better than median success.

3. Power House. Only eight boys have been assigned to the boiler tending duties of this detail. And for these boys information was given to the engineer as to their mental equipment. His ratings may not be as valid as those of the other instructors who

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 327

possessed no data except their own observation of the mental abilities of the boys. For these eight cases the Footrule correlation is found to be $.247 \pm .224$.

4. **Hospital Orderlies.** The boys in this group do the routine work of keeping the hospital clean and running errands for the nurse and patients. Only seven cases are available for inspection, and for the few cases a positive coefficient with a high Probable Error was found, $.414 \pm .211$. Little dependence can be placed upon this index except as possibly indicating a probable tendency.

SERVICE DETAILS

Under this heading are placed two details whose work is not considered as preparation for any future specific occupational employment but which may be thought of as definitely encouraging habits of industry.

1. **Housekeeping.** Under the supervision of women, small groups of boys do the janitor work in the buildings and living accommodations on the school grounds. Their duties include sweeping, scrubbing, dusting, window washing, and bed-making. This is essential routine work and is ordinarily assigned to new boys who are later detailed to more specific vocational training.

A slight negative correlation, $-.147 \pm .107$, is found between success ratings and mental age levels; a similar coefficient represents the relation of chronological age to ability to do the work. The older and brighter boys find this type of work irksome. A lower critical level of 9:0 is found, while at the same time only one out of five having mental ages above 14:6 are rated as successful.

2. **Dining Room.** This detail, under close supervision, provides the service of the meals in the large central dining room for boys, cares for the cleanliness of the room, for the setting of the tables and for the dishwashing. This detail is probationary for prospective kitchen boys and for others waiting for assignments to trade groups. There is no correlation between the success ratings and the intelligence levels, coefficient being $.034 \pm .114$. This distribution is chosen as representative of the neutral relationship between intelligence and Trade Averages.

SUMMARY AND CONCLUSIONS

1. From the consideration of the relation of trade learning records and the mental levels of those taking training in various occupations, the trades may be classified into three divisions on the basis of the nature of the relationship found to exist. The coefficients of correlation illustrate the three classifications as follows:

a. The **positive type**, in which there is found a positive coefficient, interpreted as meaning that the expected degree and oc-

currence of success is directly dependent on the position of the mental level of the boy on the scale of intelligence ratings. The positive coefficients are not large but sufficiently so to be definite indications that general intelligence is a factor in determining extent of success in learning fundamental trade processes. However, there are other factors of importance which influence the achievement.

b. The neutral type, in which the coefficient indicates a negligible relationship between mental level and potential ability to make good under the conditions described.

c. The negative type, in which the relationship is the reverse of the positive and indicates that the work is more adapted to those of lower levels of intelligence. Here again the size of the coefficients indicates the strong influence of other factors.

2. In accordance with the above divisions the trades groups studied may be grouped as follows:

a. Positive Type.

Office Boys	.983±.009	Blacksmith Shop	.299±.134
Poultry Dept.	.600±.092	Paint Shop	.283±.138
Hospital Boys414±.211	Garage	.259±.16
Print Shop	.331±.084	Power House	.247±.224
Carpenter Shop	.306±.109	Tailor Shop	.232±.072

b. Neutral Type.

Kitchen	.093±.120	Teamsters	-.087±.137
Dining Room	.030±.114	Dairy	-.126±.118
Bake Shop	-.016±.132	Housekeepers	-.147±.107
Flower Garden	-.071±.159		

c. Negative Type.

Vegetable Garden	-.234±.143	Shoe Shop	-.313±.103
Laundry	-.297±.107	Plumbing Shop	-.377±.13
Book-bindery	-.305±.14		

3. Certain of the trades show positive requirements of chronological age. But partial correlations between success and mental standing eliminating the factor of actual age failed to change materially the coefficients of relationship.

4. Minimum mental requirements for a reasonable amount of success were located as follows:

Power House	13:6	Paint Shop	11:0
Office Boys	13:0	Poultry Dept.	10:6
Blacksmith Shop	12:6	Housekeepers	9:0
Print Shop	12:3		

5. In other groups successful learning was found within the limits of the range represented and the minimum was not reached, altho the lower levels of these trades were as follows:

Shoe Shop	6:6	Teamsters	8:6
Dining Room	7:6	Vegetable Garden	8:6
Kitchen	7:6	Flower Garden	9:6
Tailor Shop	7:6	Bake Shop	10:0
Dairy	8:0	Book-bindery	10:0
Laundry	8:0	Carpenter Shop	10:0
Plumbing Shop	8:6	Garage	11:6

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 329

These findings are important as verifying on a statistical basis the fact that persons of limited mental qualifications can take part in the constructive, useful work of society and, in the case of the "negative type" occupations, with more efficiency than those of higher intelligence, provided proper supervision is given. 6. Four groups gave reversed critical levels such that 75% or more of failures occurred at mental levels higher than those indicated: Shoe Shop 12:3, Kitchen 13:3, Housekeepers 14:6, and Laundry 15:0.

7. Median or better Trade Averages in 75% or more of cases were found as follows:

Print Shop	above 15:0	Paint Shop	above 14:0
Dairy	above 14:6	Office Boys	above 13:0
Book-bindery	above 14:0	Poultry Dept.	above 12:6
Dining Room	above 14:0		

Shoe Shop	between 6:6 and 10:0
Kitchen	between 7:6 and 9:6
Laundry	between 8:0 and 10:6
Book-bindery	between 10:0 and 12:0

8. The general conclusions are that

- a. general intelligence is a factor which influences trade success during the learning stages in different degrees, depending on the trade and the conditions of instruction;
- b. in no group is the degree or fact of success wholly dependent upon intelligence;
- c. lack of a certain degree of intelligence may prevent attainment of a given degree of success;
- d. in a few trades and occupational groups statistical treatment of data indicates that general intelligence is practically negligible;
- e. the use of general intelligence measures as an aid to specific trade learning assignments is justified if sufficient care is taken in the interpretation of these ratings and if the relative value of other factors is also considered; and
- f. there is need for careful study and analysis of the other factors which assist general intelligence in the determination of potential trade ability.

REFERENCES.

1. Anderson, Miss M. L.: Education of defectives in the Public Schools. World Book Co., Yonkers, 1917.
2. Bridges and Coler: The Relation of Intelligence to Social Status. *Psych. Rev.*, XXIV-1, Jan. 1917, pp. 1-31.
3. Burt, H. E.: Prof. Münsterberg's Vocational Tests. *J. Appl. Psych.*, I, 1917, pp. 201-13.
4. Chapman, J. Crosby: Trade Tests. N. Y., Henry Holt, 1921. pp. 435.

5. Fitts, Ada M.: Mental Defectives in Industry. Ungraded, IV-8, May 1919, pp. 179-83.
6. Flanders, J. K.: Mental Tests of a Group of Employed Men, Showing Correlations with Estimates Furnished by Employer. *J. Appl. Psych.*, II, 1918. pp. 197-206.
7. Goddard, H. H.: Feeble-Mindedness: Its Causes and Consequences. p. 581.
8. Henmon, V. A.C.: Air Service Tests for Aptitude for Flying. *J. Appl. Psych.*, III-2, June 1919, pp. 103-9.
9. Kelly, R. W.: Hiring the Worker. N. Y., Eng. Mag. Co., 1918.
10. Link, H. C.: Employment Psychology. Macmillan, N. Y., 1919, pp. 440.
11. Ordahl, George: Industrial Efficiency of the Moron. *Tr. Schl. Bull.*; XV, 1919, pp. 145-53.
12. Rossy, C. S.: Feeble-Mindedness and Industrial Relations. *Mental Hygiene*, II-1, Jan. 1918. pp. 34-52.
13. Terman, L. M. and Others: A Trial of Mental and Pedagogical Tests in a Civil Service Examination for Policemen and Firemen. *J. Appl. Psych.*, I-1, Mar. 1917, pp. 17-29.
14. Thurstone, L. L.: A Standardized Test for Office Clerks. *J. Appl. Psych.*, 1919, pp. 248-51.
15. U. S. War Dept.: The Personnel System. C. C. P. N. 399. Vol. I, History of the Personnel System. C. C. P. No. 399. Vol. II, The Personnel Manual, C. C. P. No. 400.
16. Viteles, M. S.: Tests in Industry. *J. Appl. Psych.*, V-1, March 1921, pp. 57-63.
17. Woolley, Helen T.: The Mind of a Boy: Survey, XXXVII-5, Nov. 4, 1916. pp. 122-5.
18. Woolley, Helen T.: A New Scale of Mental and Physical Measurements for Adolescents and Some of its Uses. *J. Ed. Psych.*, XI-9, Nov. 1915, pp. 521-50.

MENTAL FATIGUE OF MIXED AND FULL BLOOD INDIANS.

BY THOMAS R. GARTH
University of Denver

A group of Indians composed of mixed and full bloods showed a tendency to more successfully resist the onset of mental fatigue than did whites or negroes of about the same education and age. The questions now presenting themselves are as follows:

1. Do full bloods of these groups fall away in this performance less than mixed bloods?
2. Do mixed bloods fall away less than whites?
3. As the continuous performance proceeds, what is the status of the several groups at any certain point for the curves of work when brought into relation; i. e., do any of these groups tend to hasten through sooner than others as indicated by per cent of total work done?
4. Incidentally, is the Indian's tending to resistance of the onset of fatigue here found accompanied by any conscious state indicative of this resistance, granting that the introspections obtained have any degree of reliability in such young and inexperienced subjects?

Having these questions in mind, we proceed to examine some of the data referred to above and reported by the writer in the *Journal of Applied Psychology*, Vol. IV, pp. 235-244 and Vol. V, pp. 14-25 under the titles, "Racial Differences in Mental Fatigue" and "White, Indian and Negro Work Curves."

In these articles we have already described the subjects to some extent, explaining that the Indians were of certain tribes, mostly Cherokees, Creeks, Cheyennes, and Choctaws. In fact we may safely class them as Plains and Forest Indians in contradistinction to the Plateau Indians living farther to the southwest whose origin is somewhat different. Educators working with Indians of various tribes claim to find great differences in mental capacities among the different tribes. For these and other reasons the groups forming subjects for this experiment cannot be taken as representing all the American Indian racial stocks.

There were in the original Indian groups 106 full bloods and eighty mixed bloods who were given a continuous mental task running through twenty-eight minutes for the younger sub-groups

and forty-two minutes for the older sub-groups. When we re-group these according to blood, we have in the full blood older sub-groups 47 subjects and in the younger, 61 subjects, while of the mixed bloods there are for older ones 49 subjects, and for the younger, 30. The twenty-eight and forty-two minute periods were broken up into periods of two minutes each. The records were so made that a measure of each individual's performance for any two minutes was easily obtained. These were totalled; his performance for any period was then expressed as a per cent of his total performance in the total working time. This gives an individual work curve in per cents. (See Garth, Mental Fatigue During the Continuous Exercise of a Single Function, Archives of Psychology, No. 41.) The individual work curves were then averaged, giving curves for full bloods of older Indians and of younger Indians, and

TABLE I

SHOWING ATTEMPTED PERFORMANCE

No. of Cases	3rd and 4th Grades						7th and 8th Grades					
	Mixed Bloods			Full Bloods			Mixed Bloods			Full Bloods		
Period	Av%	AD	PE	Av%	AD	PE	Av%	AD	PE	Av%	AD	PE
1	8.2	1.9	.3	6.9	1.5	.2	5.0	.5	.06	4.8	.7	.1
2	6.9	.9	.1	6.9	1.3	.1	4.5	.5	.06	4.6	.7	.1
3	7.4	1.0	.1	6.9	1.2	.1	4.5	.5	.06	4.5	.5	.07
4	7.0	.8	.1	7.1	1.5	.2	4.8	.6	.07	5.3	1.1	.2
5	7.0	1.4	.2	6.5	1.4	.2	4.6	.5	.06	4.4	.7	.1
6	7.9	1.0	.1	7.4	1.2	.1	4.7	.3	.03	4.4	1.0	.1
7	7.4	1.5	.2	6.2	1.3	.1	4.4	.6	.07	4.4	.5	.07
8	7.0	.4	.1	7.2	1.3	.1	4.8	.4	.05	5.0	.6	.08
9	6.7	.8	.1	6.3	1.2	.1	4.5	.5	.06	4.4	.6	.08
10	7.2	.3	.05	7.7	1.5	.2	4.5	.6	.07	4.7	.5	.07
11	7.2	.4	.1	7.9	1.5	.2	4.4	.4	.05	4.6	.6	.08
12	7.2	1.7	.3	7.1	1.4	.2	4.5	.2	.02	4.5	.7	.1
13	7.4	1.3	.2	7.8	1.5	.2	4.7	.4	.05	4.7	.2	.03
14	7.1	1.7	.3	7.2	.8	.2	4.2	.4	.05	4.5	.7	.1
15							4.6	.5	.06	4.7	.6	.08
16							4.2	.5	.06	4.4	.6	.08
17							4.5	.8	.1	4.8	.6	.08
18							4.6	.3	.03	5.2	.6	.08
19							4.6	.9	.1	5.0	.3	.04
20							4.7	.4	.05	5.0	.6	.09
21							4.7	.5	.06	4.9	.7	.09
Total	53.3	5.1	.9	50.9	15.2	1.7	118.0	22.8	2.8	113.0	23.5	3.4
First 3	22.6			21.0			14.1			14.4		
Last 3	22.0			22.1			14.0			14.8		
First ½	51.9			48.2			51.2			51.3		
Age 13 ⁷ / ₁₀ yrs.	2.2	.9		15.7	1.2	.9	16.8	1.6	.8	17.3	2.5	1.0
First 3												
Last 3 =	.977			1.068			.99			1.059		
	(or 2.3% loss)			(or 6.8% gain)			(or 1% loss)			(or 5.9% gain)		

for mixed bloods of older and of younger subjects, making curves for the four sub-groups as indicated above. There were two categories, the attempted performance—viz., per cents of total columns of addition attempted,—and accurate performance—being per cents of total columns actually right, which gives two curves for each blood sub-group—an attempts curve and an accurate curve. See Tables I and II.

TABLE II

No. of Cases	SHOWING ACCURATE PERFORMANCE											
	3rd and 4th Grades						7th and 8th Grades					
	Mixed Bloods			Full Bloods			Mixed Bloods			Full Bloods		
	31			61			51			47		
Period	Av%	AD	PE	Av%	AD	PE	Av%	AD	PE	Av%	AD	PE
1	9.1	4.1	.6	6.7	4.0	.4	5.3	2.2		5.2	2.4	.4
2	5.1	2.2	.3	6.5	3.7	.4	4.6	1.9		4.6	2.8	.5
3	9.1	3.8	.6	7.4	2.8	.3	5.1	1.2		6.5	2.4	.4
4	8.3	4.8	.7	7.1	3.7	.4	5.6	1.3		6.2	3.1	.5
5	7.4	3.7	.6	9.1	6.0	.7	5.8	.3		5.4	1.9	.3
6	7.1	2.0	.3	6.7	2.7	.3	5.3	1.9		5.7	2.4	.4
7	5.8	3.6	.6	6.6	4.3	.5	4.8	1.2		4.6	2.1	.3
8	3.8	.5	.1	9.2	4.8	.5	5.5	1.8		5.6	2.6	.4
9	5.8	1.4	.2	5.2	3.0	.3	3.9	1.1		4.7	2.3	.4
10	8.1	3.3	.5	7.6	4.4	.4	5.2	1.5		6.5	2.5	.4
11	7.0	3.6	.6	8.0	4.3	.4	4.6	1.5		5.2	2.3	.3
12	6.4	2.6	.4	7.0	4.0	.4	4.6	1.8		4.8	2.1	.3
13	7.6	3.2	.5	7.3	2.8	.3	4.9	1.5		6.2	2.7	.4
14	5.9	2.6	.4	6.2	3.3	.4	4.2	1.6		4.2	1.8	.3
15							4.2	1.8		5.3	2.8	.5
16							3.5	1.2		4.0	2.2	.3
17							5.4	1.6		6.1	2.9	.4
18							4.1	1.1		5.4	2.4	.4
19							5.5	1.6		5.0	2.2	.3
20							4.8	1.7		5.7	2.8	.3
21							4.4	1.7		5.5	2.8	.5
Total	28.1	10.6	1.7	29.0	13.4	1.47	69.4	22.4	9.7	64.1	18.4	3.4
First 3	23.6			20.6			15.1			14.4		
Last 3	19.9	..		22.3	..		14.7	..		14.5	..	
First ½	55.2	..		48.5	..		55.6	..		53.4	..	
First 3												
Last 3												
=												
	.81	..		1.083	..		.976	..		1.008	..	
(or 19% loss) (or 8% gain) (or 2.4% loss) (or 3% gain)												

The tables presenting the Indian data for the categories and sub-groups just described are given herewith and are composed of averages and measures of variability for any group's per cents for any successive two minutes.

While the tables give all the facts just mentioned, we have thought it well to derive curves from same in order to supply a graphic representation of these facts. In addition, we have taken

the curves for whites in the above mentioned articles (Ibid.) and placed them beside the full and mixed blood Indian curves. The curves are smoothed out by averaging the performance of each four successive minutes. Despite the small number of cases, the curves are fairly well smoothed out excepting for the younger mixed blood sub-group in the accurate performance.

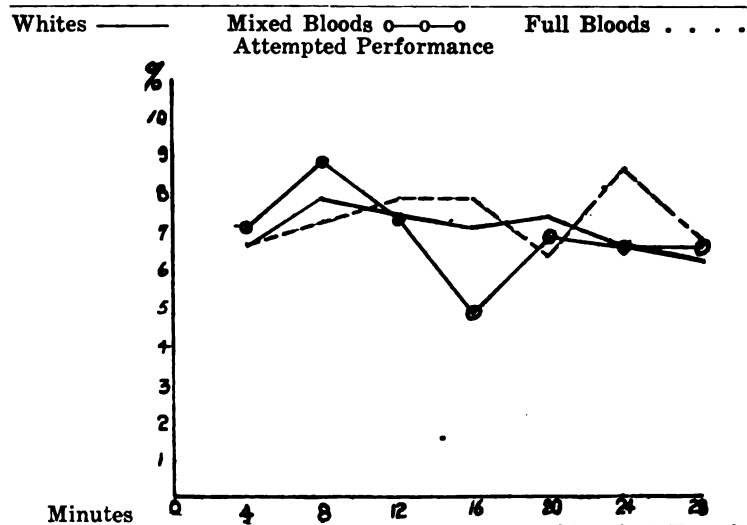


Fig. 1. Average Of The Individual Curves For Third And Fourth Grade Subjects.

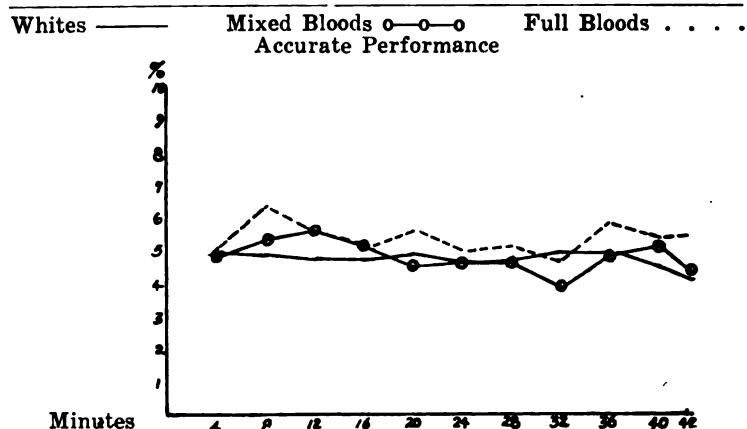


Fig. 2. Average Of The Individual Curves For Third And Fourth Grade Subjects.

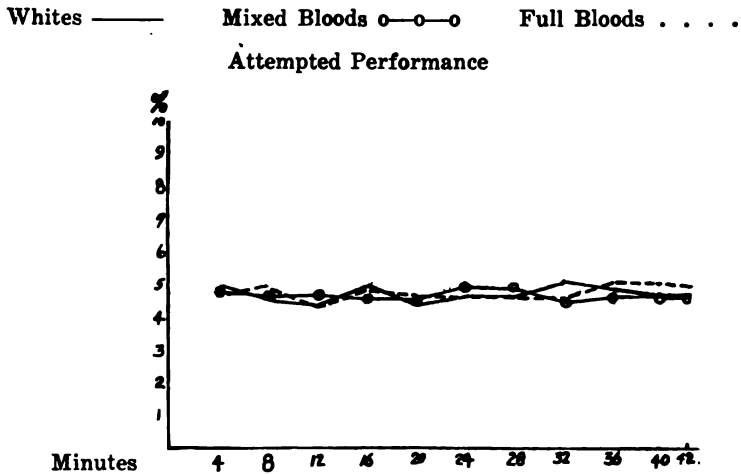


Fig. 3. Average Of The Individual Curves For The Seventh And Eighth Grade Subjects

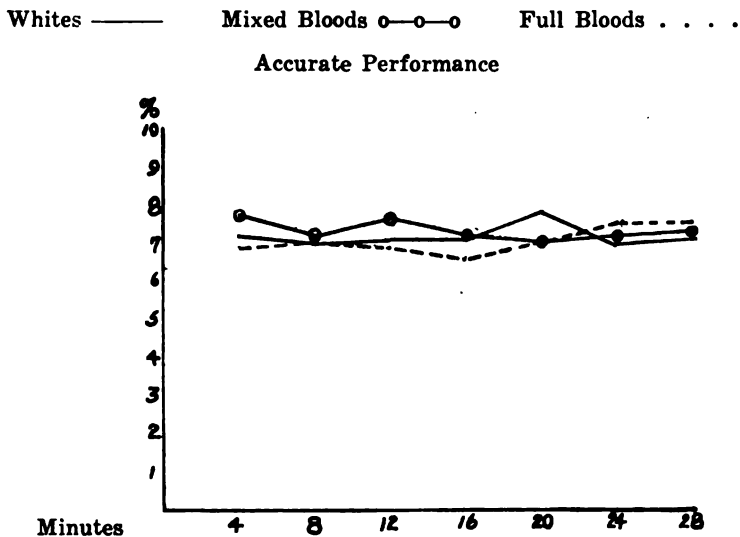


Fig. 4. Average Of The Individual Curves For The Seventh And Eighth Grade Subjects

In all the curves the Indian subjects hold up better than the whites at the end with the exception that the curves for attempted performance for the mixed bloods and the older sub-group occupy about the same points. With this exception we find the full bloods at the end holding out in the curves better than the whites, and the mixed bloods occupying a position between them.

In order to obtain some measure of the falling away of a group in its performance, the same method was used here as was used in the work by the writer mentioned above (*Archives of Psychology*, No. 41, p. 18) i. e., comparing the amount done in the first six minutes with that accomplished in the last six minutes of work.

These numbers were for the younger sub-groups:

In Attempts			In Accurates.		
Full Blood Indians	gained	6.8%	Full Blood Indians	gained	8.0%
Mixed "	"	lost 2.3%	Mixed "	"	lost 19.0%
Whites		lost 3.0%	Whites		lost 17.0%

The comparative extreme loss of the mixed bloods in accurates is due to their initial spurt which may be noted in the curve. It is possible this would not appear if a larger group of mixed blood subjects had not been obtainable. We have already called attention to the fluctuations in this curve.

For the older subjects these numbers are:

In Attempts			In Accurates		
Full Blood Indians	gained	5.9%	Full Blood Indians	gained	0.3%
Mixed "	"	lost 1.0%	Mixed "	"	lost 2.4%
Whites		lost 4.0%	Whites		lost 10.2%

This comparison favors the full bloods most and the whites least, the mixed bloods occupying an intermediate position in respect to a falling away in performance.

We call attention to the fact that the average age of the mixed bloods was, for lower grades, 13.7 years; for upper grades, 16.8 years; for full bloods it was, respectively, 15.7 years, and 17.3 years. For the whites engaged in a similar experiment the ages for the two grade sub-groups were on the average, 9.9 years and 14.05 years. It might be objected that the disparity in ages could account for the differences in fatigue, but the writer found for the data of an age for age comparison through the groups were quite small, of whites and Indians, the tendency was still in favor of the Indians (See *Racial Differences in Mental Fatigue*, Garth, *Journal of Applied Psychology*, 1919, Vol. IV).

The differences brought out by these comparisons of the curves and the falling away in performance as measured by comparing the first and last six minutes of work are not sufficient evidence of tendencies to differ in the group measures, but it will be necessary

to measure the overlapping of one distribution on another before these differences as given here are truly and significantly evaluated. While the overlapping of these "fatigue" measures cannot be secured here, we do what will probably prove just as much worth while and in doing so we bring out some facts thus far not noted. To do this we tap the curves at various points. Let the question be, in the first place, "Are the full and mixed blood curves similar or different?" In order to answer this question, we must seek the distributions whose averages compose the curves and superimpose one of the two to be compared on the other and thereby determine the amount of overlapping.

But there are too many distributions to make this a practical undertaking, so instead of doing that we tap the curves (See Journal of Applied Psychology, Vol. V, pp. 14-25) say at these points: at the beginning, at the middle, and at the end and obtain distributions for per cent of work done in first six minutes, in first half of total working period, and in last six minutes. The distributions are not given here, but only the medians and the numbers indicating the per cent of mixed bloods attaining and exceeding the

TABLE IIIA

SHOWING YOUNGER INDIAN SUBJECTS, THIRD AND FOURTH GRADES

	First 6 Minutes				First 14 Minutes				Last 6 Minutes			
	Attempts Accu.				Attempts Accu.				Attempts Accu.			
	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.
No Cases	61	30	58	31	61	30	58	30	61	31	58	31
Med.	20.5	21.3	19.3	23.0	48.3	48.5	49.5	50.0	22.5	22.0	20.3	19.8
% of M. B.												
Att. Med												
of F. B.		51%		71%		53%		63%		54%		47%
Ave.	21.4	22.6	20.6	23.6	48.2	51.9	48.5	55.2	22.1	22.0	22.3	19.5
A. D.	2.8	3.9	3.1	9.1	3.6	7.1	3.3	9.2	3.2	4.0	8.8	8.2
P. E.	2.0	3.1	2.4	7.2	2.7	6.3	2.5	7.6	2.5	3.2	7.0	6.9
									Gain	Loss	Gain	Loss
									6.8%	2.3%	8.0%	19.0%

full blood median in any distribution. Tables (III) A and B give the numbers. Table (IV) gives the white median, obtained in a previous experiment (Ibid.) and the overlapping of the median of the Indian sub-groups on the several white distributions.

In the first six minutes, then, in the case of the Indians, the mixed bloods excel the full bloods in medians and averages, and the overlapping is in their favor in all cases. In the first half of the performance (work of first fourteen minutes for the younger subjects and twenty-two minutes for the older) these measures indicate that the mixed bloods were still doing more of their work at this stage of the performance than the full bloods, with the ex-

TABLE IIIB

SHOWING OLDER INDIAN SUBJECTS, SEVENTH AND EIGHTH GRADES

	First 6 Minutes				First 22 Minutes				Last 6 Minutes			
	Attempts Accu.				Attempts Accu.				Attempts Accu.			
	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.
No. Cases	46	47	48	49	47	49	47	49	47	49	47	49
Med.	14.4	14.8	14.0	15.8	51.4	51.0	53.0	5.8	15.2	14.2	14.8	13.8
% of M. B.												
Att. Med.												
of F. B.		61%		59%		49%		59%		33%		42%
Ave.	14.0	14.1	14.4	15.1	51.3	51.2	53.4	55.6	14.8	14.0	15.4	14.7
A. D.	1.5	1.4	3.0	3.3	2.7	2.5	5.7	5.6	1.5	1.4	3.0	4.1
P. E.	1.2	1.1	2.4	2.6	2.2	2.0	4.6	4.5	1.2	1.1	2.4	3.3
									Gain	Loss	Gain	Loss
									5.9%	1.0%	.3%	2.4%

TABLE IV

SHOWING CENTRAL TENDENCIES FOR WHITES AND THE INDIAN OVERLAPPING

		Per Cent of Work Done in:					
Whites		First 6 Minutes		First 14 Minutes		Last 3 Minutes	
3rd and 4 th Grades		Att.	Acc.	Att.	Acc.	Att.	Acc.
No. Cases 368		Med. 21.81	22.41	50.28	51.52	21.59	19.82
Indians							
3rd and 4th Grades							
% Attaining		M. B.	50%	44%	50%	50%	68.1% 54.8%
Med. of Whites		F. B.	47%	28.5%	39.6%	38.6%	58.5% 70.4%
Whites		First 6 Minutes		First 22 Minutes		Last 3 Minutes	
3rd and 4th Grade		Att.	Acc.	Att.	Acc.	Att.	Acc.
No. Cases 343		Med. 14.9	15.1	53.02	54.7	14.9	14.1
Indians							
7th and 8th Grades							
% Attaining		M. B.	52%	50%	51%	48%	38% 50%
Med. of Whites		F. B.	35%	48.5%	31%	44%	56% 63%

ception of the medians for attempts in the older groups where the blood groups are about even. This would indicate that generally the mixed bloods were doing more of their work in the first half of the period than the full bloods. In the last six minutes these same measures: median, average, and that of overlapping, indicate that the full bloods resist the fatigue at the end better than the other blood group with the exception of those of the younger group in attempts for here the overlapping indicates that 50 per cent of the mixed bloods attained and exceeded the median of the full bloods.

We believe that these two defects, as shown in the overlapping, would be removed by testing larger numbers.

When we compare the Indian groups with the Whites—see Tables—the case is clearer for the full bloods as a result of taking medians and measures of overlapping. Here, the tendency for the whites to do more of their work at first and by the middle of the total time of working and less at the end than the full bloods is constant. But the numbers are not conclusive for the mixed bloods. (See table for overlapping on white distributions)¹. All we can say here is that the younger mixed bloods are similar to the whites in the first half of their performance, but somewhat better at the end and as for the older mixed bloods, the performance is similar in the first half in both categories and in accurate performance it is likewise similar at the end but not so in the attempted performance at this point.

The differences, then, are more definite between the full bloods and whites, and full and mixed bloods, than between the whites and mixed bloods.

The fourth question raised at the beginning of this discussion was as to whether the ability to resist the onset of fatigue as indicated by objective measures of the efficiency of the process here being examined, was indicated, in any way, by subjective states, i. e., by introspection, if the reports of these subjects may be taken as having reliability. In order to investigate this problem, the experimenter at the close of the experiment with the Indians requested the subjects to write on the back of the record booklet whether they were (1) Not tired, (2) Tired, (3) A little tired, or (4) Very tired. The results are interesting but rather confusing.

The reports range from a statement like the following, "I was tired, my arms surely did hurt," the writer of which was a full blood Indian girl whose accurate performance at the close was as good as at the beginning and who attempted more (22 per cent) at the close than she did at the beginning; to that of a quarter blood Caddo boy, who says he was not tired yet fell off 10 per cent in attempts and 76 per cent in accurates. The answers to the question as to whether the members of the group were tired may be summed up as follows:

Per Cent Attaining Median of Whites				Per Cent Attaining Median of Whites			
		First six	Last six			First six	Last six
Young	F.B.	Att: 47	39.6	58.6	Old	F.B.	Att: 35
		Acc: 28.5	38.6	70.4			31
							56
							63
	M.B.	Att: 50	50	68.6		M.B.	Att: 50
		Acc: 44	50	54.8			51
							38
							50

MIXED BLOODS

Reporting	Grades 3 and 4		Grades 7 and 8	
	No.	Per cent	No.	Per cent
Not tired	20	64	31	62
Little tired	8	26	15	30
Tired	0	0	4	8
Very tired	2	7	0	0
No report	1	3	0	0
Total	31	100	50	100

FULL BLOODS

Reporting	Grades 3 and 4		Grades 7 and 8	
	No.	Per cent	No.	Per cent
Not tired	26	44	20	42
Little tired	25	43	25	52
Tired	0	0	2	4
Very tired	5	9	1	2
No report	3	4	0	0
Total	59	100	48	100

It will be seen from looking at the above that over half (62 per cent) of the mixed bloods gave "not tired" as a statement of their "feelings" and that only 43 per cent of the full bloods gave this report. This would seem to show that as a group the former are poorer introspectionists than the latter, or that the resistance to fatigue may be present as a physiological fact or even absent, without the subject's being aware of it, for the mixed bloods as a group suffered more, fell away more, than the full bloods, i. e., in both the older and younger groups and in both attempts and accurates the full bloods gained (see above), and the mixed bloods lost. Over one-half of the mixed bloods reported "not tired" and 57 per cent of the full bloods reported "tired," and yet the former were doing a smaller part of their own work than the latter in both groups in both categories excepting in the 3rd and 4th groups (with 54 per cent overlapping favoring them—see table), if we take the median full blood performance as a measure. Now as for the full bloods if the median performer was doing 22.5 per cent of his total work and 20.8 per cent of his total work, respectively, (in each of the two categories reported) in the younger group; and 15.2 per cent and 14.8 per cent of his work in the older group, and 56 per cent and 58 per cent of the other group reported tired, when the fact is, they were doing as well and better than at the start, i. e., over 62 per cent of those in the younger, and 57 per cent of the older of them, we have to conclude that these full bloods either exaggerated their fatigue or had strenuous work keeping at the performance—putting forth more effort and getting better results than the mixed bloods.

We are here hovering around the problem of the reality of the old traditional stoicism of the Indian. Is this resistance to fatigue real stoicism, is it native, is it acquired? If their resistance was effortful and consequently conscious, it was not real stoicism, since stoicism is indifference to pleasure and pain, a stoic being an apathetic person. But it was due to effort if we take into consideration the objective measures and the frank statement of awareness of fatigue. If it is native it cannot be of the order of instinct, since it is marked by effort, but must be of a high order of native equipment. Sustained effort is said to mark the mental process of the genius. (See Woodworth, R. S., *Dynamic Psychology*, p. 132).

In conclusion, then, we may say:

These full blood Indians of Forest and Plains groups resist the onset of fatigue more successfully than the mixed bloods whose blood is of the same Indian stock.

The mixed blood performance is more like the white performance than that of the full bloods.

The latter's curves suggest an intermediate position at the end between the full bloods and whites, i. e., both whites and mixed bloods tend to hasten through their tasks sooner than the full bloods.

As introspectionists the full blood Indians are more inclined to acknowledge fatigue than the mixed bloods. This indicates a willingness to put forth effort in larger degree than was found in the mixed bloods.

INDIVIDUAL INJUSTICE AND GUESSING IN THE TRUE-FALSE EXAMINATION.

By J. CROSBY CHAPMAN
Yale University

On a relatively difficult fifty item True-False examination recently given by the author, a striking illustration of the unreliability of chance and the injustice done to the individual by too great reliance upon its neutralizing effect, was forthcoming. In spite of the directions which called for guessing on the unknown items, one of the subjects, after having done fourteen examples, called the examiner and said that these fourteen examples marked the limits of knowledge. Of these fourteen items, thirteen were correct. In order to simplify the marking, however, the subject was told to guess on the remaining items. When the paper was marked on the right minus wrong basis the final score of the subject was zero. While this case is undoubtedly very exceptional, the possibility of a subject's score changing from twelve to zero within a group in which the semi-interquartile range was 8.5 is, from the standpoint of reliability and individual justice, a most serious matter. That this result was not produced by elements in the test which were misleading, thereby tempting an ignorant person to make the incorrect response, was shown by the response of the other subjects which within the limits of chance tended to give on the difficult questions a fifty fifty ratio.

Kohs¹ in a paper entitled "High Test Scores by Sub-Average Minds," has shown the extent to which pure guessing in a test consisting of fifty items fails to produce a zero score. By expending the binomial theorem, he proves that the possibility of scoring plus or minus 10, on the Right minus wrong basis, is .084. He therefore justly calls in question the comfortable notion, expressed by McCall² in his article entitled "A New Kind of School Examination," that by the sure operation of chance guessing will eliminate itself. Whereas McCall implies that there is an extremely high probability that guessing will eliminate itself on a test of fifty items, Kohs shows that the half right and half wrong result will only occur once in every ten chances.

In spite of this objective evidence, McCall in his most interesting

¹Psych. Bulletin, January, 1920, page 3.

²Jour. Educ. Research, January, 1920, p. 33 ff.

book, entitled "How to Measure," has not corrected the false impression that guessing will correct itself on a test consisting of the number of items which may reasonably be expected in this form of examination.³ That this is not the result of merely loose expression of an idea, but is the outcome of an erroneous impression with regard to the operation of chance, is shown by the following statement, published in the original article and again republished in the book; which McCall presents to support his main thesis with reference to limited guessing:

"It is very difficult for some people to believe that such a test as has been outlined above does anything more than give the highest score to the luckiest guesser. They look with the eye of suspicion on this thing they call chance. I once tossed pennies for heads or tails, 50,000 times. The results came out 25,000 heads and 24,999 tails. Had there not been a miscount somewhere, the two would doubtless⁴ have come out exactly even."

Making due allowance for poetic licence with regard to the historicity of the experiment, the final statement is incorrect, and shows a failure to appreciate the significance of chance. No one can deny that in the tossing of 50,000 coins, 25,000 heads and 25,000 tails may result, but that they would have been certain to result, had there not been a miscount, is a mathematical fallacy. That 25,000 heads and 25,000 tails will be a more likely result than any other is correct, but that it will not be a certain result is obvious. In fact we can calculate the exact chance of an even break in the tossing of 50,000 coins. Using the formula expressing the standard deviation of the number of successes in events $\sigma = \sqrt{npq}$. In this case $n=50,000$ while $p=q=\frac{1}{2}$. σ is therefore approximately equal to 112. By taking the reciprocal of this quantity and referring to the normal probability curve, it will be found that the coins break even only once in approximately 280 times. The sureness therefore of which McCall writes is a sureness that occurs once in 280 times! That this statement, published in 1920, has gone unchallenged and been repeated by the same author in 1922 shows the need for correcting some of the current popular notions with reference to such problems.

In Koh's article dealing with the case of fifty items, pure guessing is assumed. This guessing on the complete number of items rarely if ever occurs in practice. The subject taking the test is sure usually of a considerable number of the items, and has merely to guess or seem to guess on the remainder. In the following

³"How to Measure" Macmillan, 1922, Page 123.

⁴Italics my own.

In all the curves the Indian subjects hold up better than the whites at the end with the exception that the curves for attempted performance for the mixed bloods and the older sub-group occupy about the same points. With this exception we find the full bloods at the end holding out in the curves better than the whites, and the mixed bloods occupying a position between them.

In order to obtain some measure of the falling away of a group in its performance, the same method was used here as was used in the work by the writer mentioned above (*Archives of Psychology*, No. 41, p. 18) i. e., comparing the amount done in the first six minutes with that accomplished in the last six minutes of work.

These numbers were for the younger sub-groups:

In Attempts			In Accurates.		
Full Blood Indians	gained	6.8%	Full Blood Indians	gained	8.0%
Mixed "	"	lost 2.3%	Mixed "	"	lost 19.0%
Whites		lost 3.0%	Whites		lost 17.0%

The comparative extreme loss of the mixed bloods in accurates is due to their initial spurt which may be noted in the curve. It is possible this would not appear if a larger group of mixed blood subjects had not been obtainable. We have already called attention to the fluctuations in this curve.

For the older subjects these numbers are:

In Attempts			In Accurates		
Full Blood Indians	gained	5.9%	Full Blood Indians	gained	0.3%
Mixed "	"	lost 1.0%	Mixed "	"	lost 2.4%
Whites		lost 4.0%	Whites		lost 10.2%

This comparison favors the full bloods most and the whites least, the mixed bloods occupying an intermediate position in respect to a falling away in performance.

We call attention to the fact that the average age of the mixed bloods was, for lower grades, 13.7 years; for upper grades, 16.8 years; for full bloods it was, respectively, 15.7 years, and 17.3 years. For the whites engaged in a similar experiment the ages for the two grade sub-groups were on the average, 9.9 years and 14.05 years. It might be objected that the disparity in ages could account for the differences in fatigue, but the writer found for the data of an age for age comparison through the groups were quite small, of whites and Indians, the tendency was still in favor of the Indians (See *Racial Differences in Mental Fatigue*, Garth, *Journal of Applied Psychology*, 1919, Vol. IV).

The differences brought out by these comparisons of the curves and the falling away in performance as measured by comparing the first and last six minutes of work are not sufficient evidence of tendencies to differ in the group measures, but it will be necessary

to measure the overlapping of one distribution on another before these differences as given here are truly and significantly evaluated. While the overlapping of these "fatigue" measures cannot be secured here, we do what will probably prove just as much worth while and in doing so we bring out some facts thus far not noted. To do this we tap the curves at various points. Let the question be, in the first place, "Are the full and mixed blood curves similar or different?" In order to answer this question, we must seek the distributions whose averages compose the curves and superimpose one of the two to be compared on the other and thereby determine the amount of overlapping.

But there are too many distributions to make this a practical undertaking, so instead of doing that we tap the curves (See Journal of Applied Psychology, Vol. V, pp. 14-25) say at these points: at the beginning, at the middle, and at the end and obtain distributions for per cent of work done in first six minutes, in first half of total working period, and in last six minutes. The distributions are not given here, but only the medians and the numbers indicating the per cent of mixed bloods attaining and exceeding the

TABLE IIIA

SHOWING YOUNGER INDIAN SUBJECTS, THIRD AND FOURTH GRADES

	First 6 Minutes				First 14 Minutes				Last 6 Minutes			
	Attempts Accu.				Attempts Accu.				Attempts Accu.			
	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.
No Cases	61	30	58	31	61	30	58	30	61	31	58	31
Med.	20.5	21.3	19.3	23.0	48.3	48.5	49.5	50.0	22.5	22.0	20.3	19.8
% of M. B.												
Att. Med												
of F. B.		51%		71%		53%		63%		54%		47%
Ave.	21.4	22.6	20.6	23.6	48.2	51.9	48.5	55.2	22.1	22.0	22.3	19.5
A. D.	2.8	3.9	3.1	9.1	3.6	7.1	3.3	9.2	3.2	4.0	8.8	8.2
P. E.	2.0	3.1	2.4	7.2	2.7	6.3	2.5	7.6	2.5	3.2	7.0	6.9
									Gain	Loss	Gain	Loss
									6.8%	2.3%	8.0%	19.0%

full blood median in any distribution. Tables (III) A and B give the numbers. Table (IV) gives the white median, obtained in a previous experiment (Ibid.) and the overlapping of the median of the Indian sub-groups on the several white distributions.

In the first six minutes, then, in the case of the Indians, the mixed bloods excel the full bloods in medians and averages, and the overlapping is in their favor in all cases. In the first half of the performance (work of first fourteen minutes for the younger subjects and twenty-two minutes for the older) these measures indicate that the mixed bloods were still doing more of their work at this stage of the performance than the full bloods, with the ex-

TABLE IIIB

SHOWING OLDER INDIAN SUBJECTS, SEVENTH AND EIGHTH GRADES

	First 6 Minutes				First 22 Minutes				Last 6 Minutes			
	Attempts Accu.				Attempts Accu.				Attempts Accu.			
	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.	F.B.	M.B.
No. Cases	46	47	48	49	47	49	47	49	47	49	47	49
Med.	14.4	14.8	14.0	15.8	51.4	51.0	53.0	5.8	15.2	14.2	14.8	13.8
% of M. B.												
Att. Med.												
of F. B.		61%		59%		49%		59%		33%		42%
Ave.	14.0	14.1	14.4	15.1	51.3	51.2	53.4	55.6	14.8	14.0	15.4	14.7
A. D.	1.5	1.4	3.0	3.3	2.7	2.5	5.7	5.6	1.5	1.4	3.0	4.1
P. E.	1.2	1.1	2.4	2.6	2.2	2.0	4.6	4.5	1.2	1.1	2.4	3.3
									Gain	Loss	Gain	Loss
									5.9%	1.0%	.3%	2.4%

TABLE IV

SHOWING CENTRAL TENDENCIES FOR WHITES AND THE INDIAN OVERLAPPING

		Per Cent of Work Done in:					
Whites	First 6 Minutes	First 14 Minutes		Last 3 Minutes			
3rd and 4th	Att.	Acc.	Att.	Acc.	Att.	Acc.	
Grades	Med.	21.81	22.41	50.28	51.52	21.59	19.82
No. Cases 368							
Indians							
3rd and 4th							
Grades							
% Attaining	M. B.	50%	44%	50%	50%	68.1%	54.8%
Med. of Whites	F. B.	47%	28.5%	39.6%	38.6%	58.5%	70.4%
Whites							
3rd and 4th	First 6 Minutes	First 22 Minutes		Last 3 Minutes			
Grade	Att.	Acc.	Att.	Acc.	Att.	Acc.	
No. Cases 343	Med.	14.9	15.1	53.02	54.7	14.9	14.1
Indians							
7th and 8th							
Grades							
% Attaining	M. B.	52%	50%	51%	48%	38%	50%
Med. of Whites	F. B.	35%	48.5%	31%	44%	56%	63%

ception of the medians for attempts in the older groups where the blood groups are about even. This would indicate that generally the mixed bloods were doing more of their work in the first half of the period than the full bloods. In the last six minutes these same measures: median, average, and that of overlapping, indicate that the full bloods resist the fatigue at the end better than the other blood group with the exception of those of the younger group in attempts for here the overlapping indicates that 50 per cent of the mixed bloods attained and exceeded the median of the full bloods.

We believe that these two defects, as shown in the overlapping, would be removed by testing larger numbers.

When we compare the Indian groups with the Whites—see Tables—the case is clearer for the full bloods as a result of taking medians and measures of overlapping. Here, the tendency for the whites to do more of their work at first and by the middle of the total time of working and less at the end than the full bloods is constant. But the numbers are not conclusive for the mixed bloods. (See table for overlapping on white distributions)¹. All we can say here is that the younger mixed bloods are similar to the whites in the first half of their performance, but somewhat better at the end and as for the older mixed bloods, the performance is similar in the first half in both categories and in accurate performance it is likewise similar at the end but not so in the attempted performance at this point.

The differences, then, are more definite between the full bloods and whites, and full and mixed bloods, than between the whites and mixed bloods.

The fourth question raised at the beginning of this discussion was as to whether the ability to resist the onset of fatigue as indicated by objective measures of the efficiency of the process here being examined, was indicated, in any way, by subjective states, i. e., by introspection, if the reports of these subjects may be taken as having reliability. In order to investigate this problem, the experimenter at the close of the experiment with the Indians requested the subjects to write on the back of the record booklet whether they were (1) Not tired, (2) Tired, (3) A little tired, or (4) Very tired. The results are interesting but rather confusing.

The reports range from a statement like the following, "I was tired, my arms surely did hurt," the writer of which was a full blood Indian girl whose accurate performance at the close was as good as at the beginning and who attempted more (22 per cent) at the close than she did at the beginning; to that of a quarter blood Caddo boy, who says he was not tired yet fell off 10 per cent in attempts and 76 per cent in accurates. The answers to the question as to whether the members of the group were tired may be summed up as follows:

	Per Cent Attaining Median of Whites					Per Cent Attaining Median of Whites			
	First six	First half	Last six	Last six		First six	First half	Last six	Last six
Young	F.B. Att: 47	39.6	58.6		Old	F.B. Att: 35	31	56	
	Acc: 28.5	38.6	70.4			Acc: 48.5	44	63	
	M.B. Att: 50	50	68.6			M.B. Att: 50	51	38	
	Acc: 44	50	54.8			Acc: 50	48	50	

The relation found between the ratings and the mental levels of those in the group shows the highest positive correlation, except one, of all the groups studied. The coefficient is $.60 \pm .092$.

Two clear cut critical levels are indicated at the 12:6 level and the 10:6 division.

There is a similar but not so marked relationship between actual age and success.

The most effective learning is achieved by boys of the higher ages whose mental levels approximate or exceed 12 years 6 months.

4. Vegetable Gardens. Supervised care of plant propagation and cultivation is the field of activity of this detail. All planning of planting is done by the instructor and farm manager. The boys are essentially supervised workers.

For this group a negative correlation, coefficient $-.243 \pm .143$, is found. Most satisfactory results are obtained from the boys of the lower levels of intelligence. As suggested by Rossy (12), this type of work is especially adapted to defectives.

The positive correlation .307, altho small, suggests chronological age rather than mental ability as a criterion for selecting workers in this detail.

MISCELLANEOUS UNCLASSIFIED

1. Office Boys. Altho a very small group of boys has been assigned to this work of running errands for the administrative offices of the School, the distribution of their ratings is interesting as indicating almost perfect correlation with mental level. Correlation coefficient $.983 \pm .009$. The boys who made median or better success were all of mental ages above 13 years 3 months.

2. Flower Garden. The boys of this group are assigned to routine tasks such as hoeing weeds, trimming bushes, picking flowers, but under comparatively lax supervision. The emphasis in the selection of boys placed on dependability rather than upon other factors seems to be at least partially warranted by the relation found between two other factors and trade ratings. The correlation with mental age, $-.071 \pm .150$, and that with chronological age, $-.169 \pm .146$, are both practically zero and seem to be almost negligible in selection considerations. The evidence is insufficient to validate a critical level suggested at mental age 14:6. Both of the two cases above that level attain to better than median success.

3. Power House. Only eight boys have been assigned to the boiler tending duties of this detail. And for these boys information was given to the engineer as to their mental equipment. His ratings may not be as valid as those of the other instructors who

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 327

possessed no data except their own observation of the mental abilities of the boys. For these eight cases the Footrule correlation is found to be $.247 \pm .224$.

4. **Hospital Orderlies.** The boys in this group do the routine work of keeping the hospital clean and running errands for the nurse and patients. Only seven cases are available for inspection, and for the few cases a positive coefficient with a high Probable Error was found, $.414 \pm .211$. Little dependence can be placed upon this index except as possibly indicating a probable tendency.

SERVICE DETAILS

Under this heading are placed two details whose work is not considered as preparation for any future specific occupational employment but which may be thought of as definitely encouraging habits of industry.

1. **Housekeeping.** Under the supervision of women, small groups of boys do the janitor work in the buildings and living accommodations on the school grounds. Their duties include sweeping, scrubbing, dusting, window washing, and bed-making. This is essential routine work and is ordinarily assigned to new boys who are later detailed to more specific vocational training.

A slight negative correlation, $-.147 \pm .107$, is found between success ratings and mental age levels; a similar coefficient represents the relation of chronological age to ability to do the work. The older and brighter boys find this type of work irksome. A lower critical level of 9:0 is found, while at the same time only one out of five having mental ages above 14:6 are rated as successful.

2. **Dining Room.** This detail, under close supervision, provides the service of the meals in the large central dining room for boys, cares for the cleanliness of the room, for the setting of the tables and for the dishwashing. This detail is probationary for prospective kitchen boys and for others waiting for assignments to trade groups. There is no correlation between the success ratings and the intelligence levels, coefficient being $.034 \pm .114$. This distribution is chosen as representative of the neutral relationship between intelligence and Trade Averages.

SUMMARY AND CONCLUSIONS

1. From the consideration of the relation of trade learning records and the mental levels of those taking training in various occupations, the trades may be classified into three divisions on the basis of the nature of the relationship found to exist. The coefficients of correlation illustrate the three classifications as follows:

a. The **positive type**, in which there is found a positive coefficient, interpreted as meaning that the expected degree and oc-

ticeable that of those with mental ages below 12:0, the deviations from median success of the failures are greater than those of the successes. In other words, with about even numerical proportions of successes and failures, the degree of the success of those below the median mental age is comparatively limited.

Although boys of mental ages as low as 10 years can learn, those individuals whose intelligence rates above 12:3 attain to greater degrees of success.

2. **Paint Shop.** The work of the painting detail consists of the finishing of furniture, including flat painting, staining, varnishing, and rubbed finishing; interior painting and decorating, and floor surfacing; outside painting; auto body work; curtain-hanging and glazing. The boys are given instruction in this work, as well as in the preparation of materials, color mixing, care of brushes and tools, and some training in sign construction and lettering.

The instructor is a painter of considerable trade experience and holds a special certificate for vocational instruction.

For the group of 20 boys the distribution presents a positive relationship between mental and trade ability, correlation coefficient $.283 \pm .138$. The tendency is small but definite. The diagram suggested both an upper and a lower critical level at 14:0 and 11:0 respectively. Not a single median rating was attained by any boy whose mental level was below 11 years. Four out of five above the upper level were rated at the median or better.

The negative relationship between chronological age and success suggests that boys can most profitably be started in the learning of this trade before they are 16 years of age.

The minimum mental age of 11 years is required for successful learning of the operations involved and the highest proportion of success will be expected of boys between 14 and 16 years old whose mental levels are 11 years or higher, preferably above 14 years.

3. **Plumbing Shop.** The plumbing shop duties consist of general repair work on water, gas, and steam service lines, with little new installation work. The instructor is a journeyman plumber in charge of from one to three boys at a time. He has had no special instructional training.

The limited number of boys observed renders unreliable any statistical indices. The eleven cases rated gave a negative coefficient of correlation between mental age levels and Trade Averages, $-.377 \pm .13$. No general conclusions seem to be warranted although this would seem to place this group in the class of trade where intelligence is not an important factor in the early learning process.

4. **Blacksmith and Machine Shop.** This detail handles all of the

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 321

metal working for the school, including forge work and a limited amount of machine practice in the use of lathes, shapers, milling machines and drill presses. The instructor is a forge man who has had some additional machine operation.

The coefficient $.299 \pm .134$ represents a limited positive relation between intelligence and success. A lower critical level is suggested at M. A. 12:6, but the cases below that level are barely inferior to those attaining median success. Without more cases this level is questionable. However the highest degree of success is attained by those whose mental levels are above 13:6.

Similarly a chronological age of 15:9 or higher predicts the most likelihood of efficient learning.

5. **Garage.** The work of the boys in the garage detail consists primarily in the observation of repair work and actual practice in the general maintenance and driving of automobiles and motor trucks. The detail is small and renders questionable the meaning of a positive correlation coefficient of $.259 \pm .16$.

GARMENT TRADES

1. **Tailor Shop.** The operations performed in the tailor shop come under three heads, (1) the mending and darning of worn and torn garments, (2) the sorting and inspection of laundered clothing and its issuance to the boys of the School, and (3) the making of new underwear, shirts, work trousers, caps and suits. The work is under the supervision of a woman whose experience in the trade has been largely that of institutional tailoring instruction. The detail contains an average of 20 boys.

The ratings of the instructor are contained in a comparatively narrow range of the rating scale and are therefore not as reliable as the distribution represented in the charts. ✓

For the 78 cases included in this detail a slight positive relation is found, $.232 \pm .072$. But in the wide range of mental ages, 7:6 to 17:8, no critical levels are indicated. Apparently very low intelligences are sufficient for productive work in this shop, and the lower critical level falls below 7 years 6 months.

2. **Shoe Shop.** Shoes are both made and repaired in the shoe shop. A few other leather articles such as footballs are cut and made from pattern. Each boy passes in succession thru the various processes involved in the making of the shoes, but since the styles of shoe are limited, the number and variety of problems faced are comparatively few. There is considerable repetition of activity.

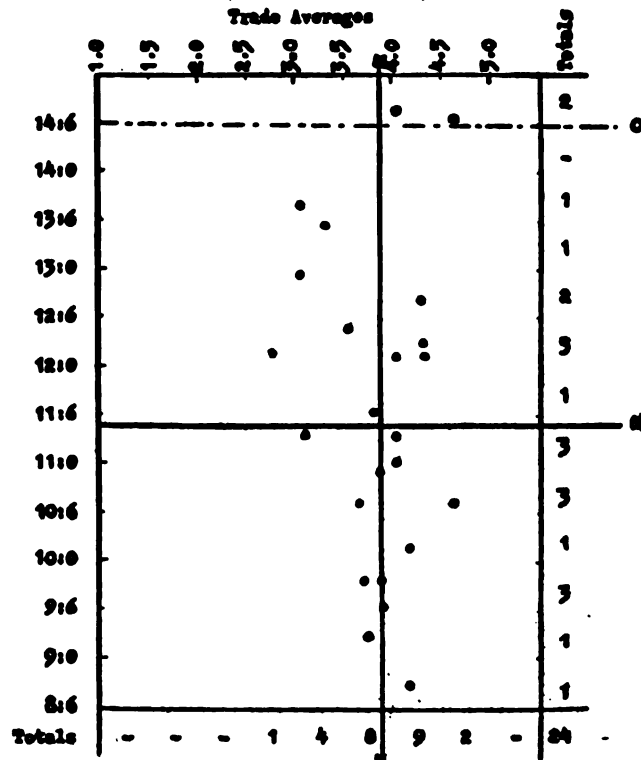
A distinct negative relationship is found between mental rating and Trade Averages, $-.313 \pm .103$. Only a very low level of intelligence is required for successful learning, the minimum level

falls somewhere below the lower limit of our range, 6:6. On the other hand the higher levels of ability find difficulty in adapting themselves to the operations and restricted supervised activities of the trade learning under the shop conditions. Chart III represents the distribution and shows the reversed critical levels at years 10:0 and 12:3.

CHART II

TEAMSTERS

Distribution of Trade Averages by Mental Age Levels
(Neutral Relation)



LEGEND

M Median

C Suggested Possible Critical Level

3. Laundry. The work of the boys in the laundry consists of hand ironing and attendance upon the comparatively simple machines in use in modern laundries. The instructor supervises much

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 323

as does the foreman in commercial work, and also cares for the up-keep and operation of the more complicated machines. Under these circumstances the group assigned to the detail has been of lower average intelligence than those working in the previously discussed groups, median mental age 11:10. The range, however, has been fairly comprehensive, 8:0 to 16:5. For the group the correlation of mental age with trade success is $-.297 \pm .107$, and inspection of the distribution revealed consistent failure for the boys whose mental ages were above 15 years. Likewise the successes are most frequent among those below the 10:6 level. The critical levels are reversed and definite. A minimum requirement of intelligence is not defined within the range represented.

CULINARY TRADES

1. Kitchen. All food except the bread and pastries made in the bake shop is prepared for serving by the kitchen. For purposes of segregation and because of the commercial opportunities after leaving the school, this detail is made up largely of negro boys.

A range of intelligence levels from 7:6 to 15:8 is included in the group, with a median mental age of 12:3. The chef failed to make use of the larger part of the rating scale, with the result that the lowest Trade Average is above the median of the ratings given by the other instructors. It is questionable whether every boy in the detail made the success indicated. However variations within the group are present altho in very slight amounts. The resulting coefficient of correlation, $.093 \pm .12$, is practically zero. There seems to be little if any relation between intelligence and comparative success. However, at each extreme of the range there is a group of which the majority fail to attain to the median rating of the group. Above mental age 13:3, not a single case reached the median rating, while a similar distribution occurs below the level 9:6. Greatest success occurs between the mental ages 9:6 and 13:3, but the lowest rating as noted above is comparatively high. ✓

2. Bake Shop. In the bake shop are prepared the bread and pastries (pies, cakes, cookies, doughnuts, etc.) for the dining rooms. The boys do the work of mixing the doughs, timing the rising and baking processes, and remove the completed articles from the ovens; all under the supervision and instruction of a first-class baker.

Here again the relation between mental rating and Trade Averages is practically zero, $.016 \pm .132$, altho the range of rating is complete. The factor of intelligence is nearly negligible under the conditions of supervision provided in this shop.

On the other hand chronological age is a factor to be used as a

guide in the selection of boys, the coefficient of correlation with Trade Averages being $.476 \pm .102$. A chronological age of 14:6 is required for reasonable expectation of success and both higher proportions and degrees of success are attained by the still older boys.

FARM WORK

1-2. Dairy and Teamsters. The similarity of the statistical indications and of the diagrammatic distributions for these two details makes possible discussion of the two groups together after the duties of each have been mentioned.

The work in the dairy consists in the care of the physical equipment, barns, stalls, milk receptacles, feed and feed-bins, of the feeding of the cows, and handling of the milk after cooling and separation. The milking is done by hired milkers. Assignments are made from the dairy detail to the feeding of the hogs and the care of their pens.

The teamsters are supervised in the care of horses, wagons, and the horse barns. They drive the teams for all hauling service not handled by motor trucks, for plowing, cultivating, etc., where tractor operation is not feasible. All tractor work is done by hired help. In two cases the tractor operators have been boys who have formerly been members of the teaming detail and who have been hired after special instruction in tractor operation.

In the case of each of these groups the relation of mental age to subsequent success in the work is very small and negative (practically zero, in view of the P. E.s). Coefficients are for the dairy $-.126 \pm .118$, and for the teamsters $-.087 \pm .137$. The distribution of Trade Averages by mental ages consist of nearly even divisions of successes and failures at all levels from mental age 8:0 to 14:6. Chart II gives the distribution of teamster ratings and illustrates the mental relationships. In each group all boys with mental ages above 14:6 receive better than median averages. This suggests a possible upper critical Level at this point. Likewise the relation between chronological age and success is zero.

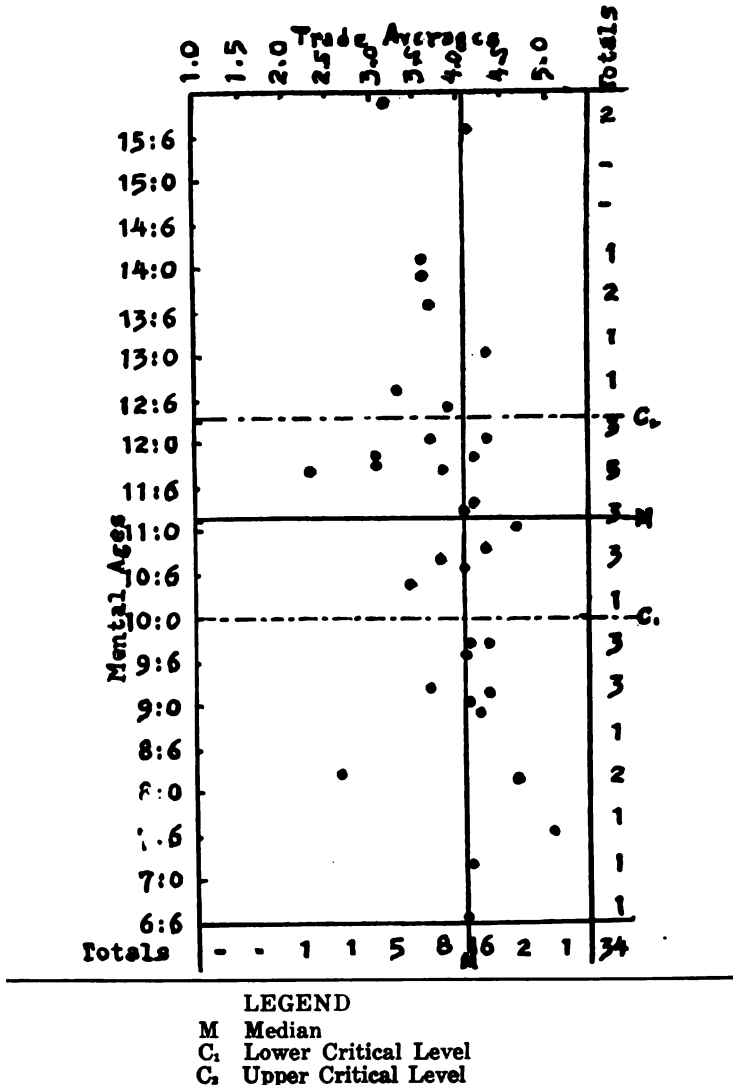
The practical suggestion is that boys between ages 13 and 19 and whose mental ages are above 14:6 can be expected to make successful progress in this work; and that for mental ages below 14:6 the expectation is about equally divided between success and failure. The minimum intelligence required is below the level of 8 years.

3. Poultry Department. Instruction is given in the feeding of poultry, breeding of pure strain stock, care of pens, handling of egg-laying and records, incubation and hatching. The instructor is an experienced chicken fancier and egg producer, but has had no training in educational methods.

CHART III

SHOE SHOP

Distribution of Trade Averages by Mental Age Levels
(Negative Relation)



The relation found between the ratings and the mental levels of those in the group shows the highest positive correlation, except one, of all the groups studied. The coefficient is $.60 \pm .092$.

Two clear cut critical levels are indicated at the 12:6 level and the 10:6 division.

There is a similar but not so marked relationship between actual age and success.

The most effective learning is achieved by boys of the higher ages whose mental levels approximate or exceed 12 years 6 months.

4. Vegetable Gardens. Supervised care of plant propagation and cultivation is the field of activity of this detail. All planning of planting is done by the instructor and farm manager. The boys are essentially supervised workers.

For this group a negative correlation, coefficient $-.243 \pm .143$, is found. Most satisfactory results are obtained from the boys of the lower levels of intelligence. As suggested by Rossy (12), this type of work is especially adapted to defectives.

The positive correlation .307, altho small, suggests chronological age rather than mental ability as a criterion for selecting workers in this detail.

MISCELLANEOUS UNCLASSIFIED

1. Office Boys. Altho a very small group of boys has been assigned to this work of running errands for the administrative offices of the School, the distribution of their ratings is interesting as indicating almost perfect correlation with mental level. Correlation coefficient $.983 \pm .009$. The boys who made median or better success were all of mental ages above 13 years 3 months.

2. Flower Garden. The boys of this group are assigned to routine tasks such as hoeing weeds, trimming bushes, picking flowers, but under comparatively lax supervision. The emphasis in the selection of boys placed on dependability rather than upon other factors seems to be at least partially warranted by the relation found between two other factors and trade ratings. The correlation with mental age, $-.071 \pm .150$, and that with chronological age, $-.169 \pm .146$, are both practically zero and seem to be almost negligible in selection considerations. The evidence is insufficient to validate a critical level suggested at mental age 14:6. Both of the two cases above that level attain to better than median success.

3. Power House. Only eight boys have been assigned to the boiler tending duties of this detail. And for these boys information was given to the engineer as to their mental equipment. His ratings may not be as valid as those of the other instructors who

INTELLIGENCE AND SUCCESS IN TRADE LEARNING 327

possessed no data except their own observation of the mental abilities of the boys. For these eight cases the Footrule correlation is found to be $.247 \pm .224$.

4. **Hospital Orderlies.** The boys in this group do the routine work of keeping the hospital clean and running errands for the nurse and patients. Only seven cases are available for inspection, and for the few cases a positive coefficient with a high Probable Error was found, $.414 \pm .211$. Little dependence can be placed upon this index except as possibly indicating a probable tendency.

SERVICE DETAILS

Under this heading are placed two details whose work is not considered as preparation for any future specific occupational employment but which may be thought of as definitely encouraging habits of industry.

1. **Housekeeping.** Under the supervision of women, small groups of boys do the janitor work in the buildings and living accommodations on the school grounds. Their duties include sweeping, scrubbing, dusting, window washing, and bed-making. This is essential routine work and is ordinarily assigned to new boys who are later detailed to more specific vocational training.

A slight negative correlation, $-.147 \pm .107$, is found between success ratings and mental age levels; a similar coefficient represents the relation of chronological age to ability to do the work. The older and brighter boys find this type of work irksome. A lower critical level of 9:0 is found, while at the same time only one out of five having mental ages above 14:6 are rated as successful.

2. **Dining Room.** This detail, under close supervision, provides the service of the meals in the large central dining room for boys, cares for the cleanliness of the room, for the setting of the tables and for the dishwashing. This detail is probationary for prospective kitchen boys and for others waiting for assignments to trade groups. There is no correlation between the success ratings and the intelligence levels, coefficient being $.034 \pm .114$. This distribution is chosen as representative of the neutral relationship between intelligence and Trade Averages.

SUMMARY AND CONCLUSIONS

1. From the consideration of the relation of trade learning records and the mental levels of those taking training in various occupations, the trades may be classified into three divisions on the basis of the nature of the relationship found to exist. The coefficients of correlation illustrate the three classifications as follows:

a. The **positive type**, in which there is found a positive coefficient, interpreted as meaning that the expected degree and oc-

TABLE IX PERCENTILE SCORES OF THE TWENTY-FIVE COLLEGE SUBJECTS

Tests	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
Army Intelligence	40	27	1	15	60	40	94	62	99	20	80	68	5	46	17	85	30	60	72	28	34	50	80	93	72
Thurstone Intelligence	35	40	7	17	35	45	95	55	99	26	77	5	62	80	1	80	60	60	83	92	76	22	45	30	52
Binet Mental Age	60	70	1	10	30	40	80	40	99	8	20	80	90	35	99	34	90	70	55	30	60	15	55	60	
Binet I. Q.	60	75	1	15	30	45	80	55	99	10	20	80	80	90	40	99	34	90	70	59	20	60	15	86	60
Completion	50	65	4	14	1	35	99	20	99	12	40	54	40	45	28	84	25	90	70	80	50	75	35	90	76
Opposites	1	15	70	40	25	75	95	1	99	40	17	13	65	80	80	48	34	60	35	85	60	85	65	55	35
Reading	60	56	32	15	8	17	92	20	99	1	55	72	42	74	25	85	20	45	85	80	40	33	70	93	78
Range of Interest	28	16	20	1	60	50	95	80	99	80	50	70	10	25	77	80	40	50	36	30	70	50	8	95	70
Free Association	15	60	1	65	9	34	68	80	99	46	48	50	93	88	77	55	99	20	24	40	72	28	90	32	16
Linguistic Invention	40	60	8	1	4	48	80	60	95	17	20	48	35	99	94	80	12	70	35	53	80	80	66	20	70
Word Building	40	77	1	8	30	15	80	25	60	60	40	65	85	92	20	99	92	36	50	50	70	16	90	70	30
Visual App. (Tachis.)	65	50	30	30	1	25	28	65	96	1	3	65	80	50	80	20	99	65	80	50	20	40	55	90	
Visual App. (Objects)	55	75	70	75	15	30	10	90	99	30	30	55	90	30	80	55	10	30	30	30	30	1	55	55	90
Aussage (Australians)	56	40	40	53	9	10	80	96	99	65	20	60	1	15	87	92	26	80	30	65	88	25	72	45	55
Aussage (Disputed Case)	85	80	34	60	1	18	92	99	60	10	40	60	20	23	35	55	50	92	70	72	88	46	25	47	14
Logical Memory	30	85	19	32	1	60	70	92	99	12	20	50	68	57	40	50	20	95	85	80	72	40	68	30	8
Rote Memory (Digits)	70	55	55	40	30	30	20	55	90	16	20	55	90	55	90	90	15	90	40	55	1	40	55	40	90
Rote Memory (Words)	90	90	35	70	20	40	40	30	40	20	10	55	99	75	70	55	15	55	36	75	10	30	55	1	55
Learning (Substitution)	55	68	42	45	12	40	8	75	99	1	20	27	60	35	20	80	32	85	60	92	72	28	80	62	89
Learning (Mirror Drawing)	48	20	1	50	7	83	60	60	75	25	30	54	86	60	80	50	60	99	96	12	30	25	99	10	40
Cancellation	23	72	15	12	77	35	67	46	84	20	60	9	99	1	82	80	58	45	60	95	93	40	35	65	10
Average	48	57	23	32	22	39	68	57	90	25	34	52	61	58	55	75	37	69	57	62	54	41	55	54	55

TESTS AND ANALYSIS OF MENTAL CONSTITUTION 361

among a hundred persons instead of by actual score. The test scores of the subjects, who are discussed in detail, are shown graphically in Figures I to IV inclusive. The base line of the probability curve of frequency distribution for which the percentile values are known, made a suitable scale for graphic representation.

The fact that the differences in score between the fiftieth and sixtieth percentile is less than between the ninetieth and one-hundredth percentile is made obvious by this method.

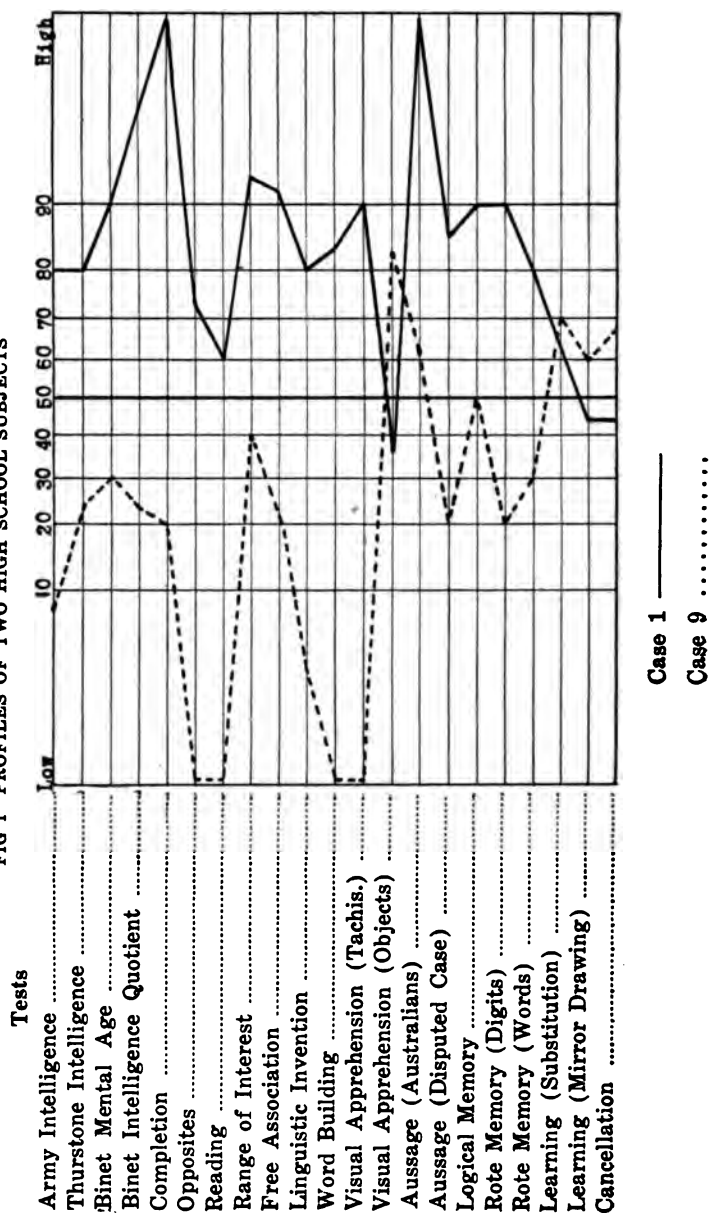
Four cases from the High-school subjects have been selected for discussion. They are Nos. 1, 9, 13, and 19 of Table VIII in the column headed "Tests."

Case No. 1. (Very superior and consistent mental ability) Boy; age 14:4; sophomore; average grade 92; mental age 18-1; I. Q. 127. He is a boy of superior ability, and although he is more than two years below the average age of the group with whom he is being compared, he is above the fiftieth percentile in all but three mental tests—three which do not show a *good* correlation with general ability. He made the highest record in the completion tests and the first "Aussage" test, and is in the vicinity of the eightieth or ninetieth percentiles in almost all of the tests. His army test score is 145, which is above the average score for the sophomores in one of the largest universities. Although he was urged to exert himself on these tests, he seemed to do all of them with ease, and showed no emotional disturbances in the most perplexing situations.

He is slightly under size for his age, dark, with a pleasing though somewhat sophisticated manner. It was difficult to find time to complete the mental tests on account of his many activities. He is carrying such a heavy schedule of school work that he has only three periods available per week for studying at school. Yet after school hours he was practicing for track, so that he had to prepare most of his lessons in the evening. He is also active in the Boy Scouts and is taking piano lessons. His mother is a talented musician. Up to a short time ago he had delivered newspapers, and in the oral "Free Association" test he named a list of city streets at an amazing rate, until it was thought wise to suggest some different topic, whereupon he gave a longer list of articles of clothing with even greater speed.

It was very evident that his relations with his parents are happy. His chief ambition, after finishing college, is to go into business with his father, who is a broker. One cannot help but feel that when his interests are thoroughly awakened he will achieve marked success in some one of the higher professions.

FIG 1 PROFILES OF TWO HIGH SCHOOL SUBJECTS



TESTS AND ANALYSIS OF MENTAL CONSTITUTION 363

This boy's school record is excellent and his teachers rate him very high. He carries his heavy schedule with ease. He is going through the high school as rapidly as possible and expects to enter college immediately thereafter. Exceptional ability linked with great industry and variety of interests, make a fortunate combination in this lad which should guarantee his success.

Case No. 9. (Inferior mental ability.) Boy; age 16-7; eighth grade; average grade 56; mental age 15-0; I. Q. 94.

There are certain striking constitutional differences in this boy's case which are not disclosed in tests given by the group method. His specific abilities are on a slightly better level than his general ability, but even in those he has peculiar records. The low record in "Reading" and "Visual Apprehension" is of most significance. Apparently he has a mild form of aphasia, since his visual apprehension of non-verbal material was nearly normal, but that for verbal material was very low. Tests involving rapid rates of reading uniformly brought low records. He wears glasses which give normal vision.

He is said to have some ability in drawing, and during the past semester he earned a grade of 8.5 in mechanical drawing. This boy attended school in Detroit for several years, where he repeated the work of the fifth grade. Since then he has always had difficulty in earning promotion. He failed in the eighth grade last year, but on account of his age and size, (he is tall and fully developed physically) he was advanced to the high school on trial. He was unable to carry the work there and was removed to the eighth grade, where he failed again in the traditional subjects.

He claims that he failed in school on account of general ill health, and also that he has a weak heart. He comes from an average family. An older sister is completing her university course this year. He is not seriously concerned about his poor progress or about his handicap in reading. Although he was urged to have a special tutor in reading, to our knowledge he has done nothing about it.

His example and presence in the classroom are very trying to his teachers; his handwriting, for instance, is practically illegible. He should be in a special room. Although he is beyond the legal school age, he shows no inclination to leave school or consider work. It is the lesser of the two evils to remain in school as long as possible. The school should do its best to fit him for a semi-skilled trade, where his ability for drawing would be utilized. At present his time and that of the teacher is practically wasted.

It was only by careful and somewhat extensive work with many mental tests that his difficulties were discovered. His mentality,

judged by a group test, would be rated low, and he would be relegated to a slow section upon a flexible plan of promotion. In his case the group method was not sufficient to determine the chief trouble.

Case No. 13. (Erratic and generally inferior mental ability.) Girl; age 17-3; freshman; average grade 67; mental age 14-8; I. Q. 89.

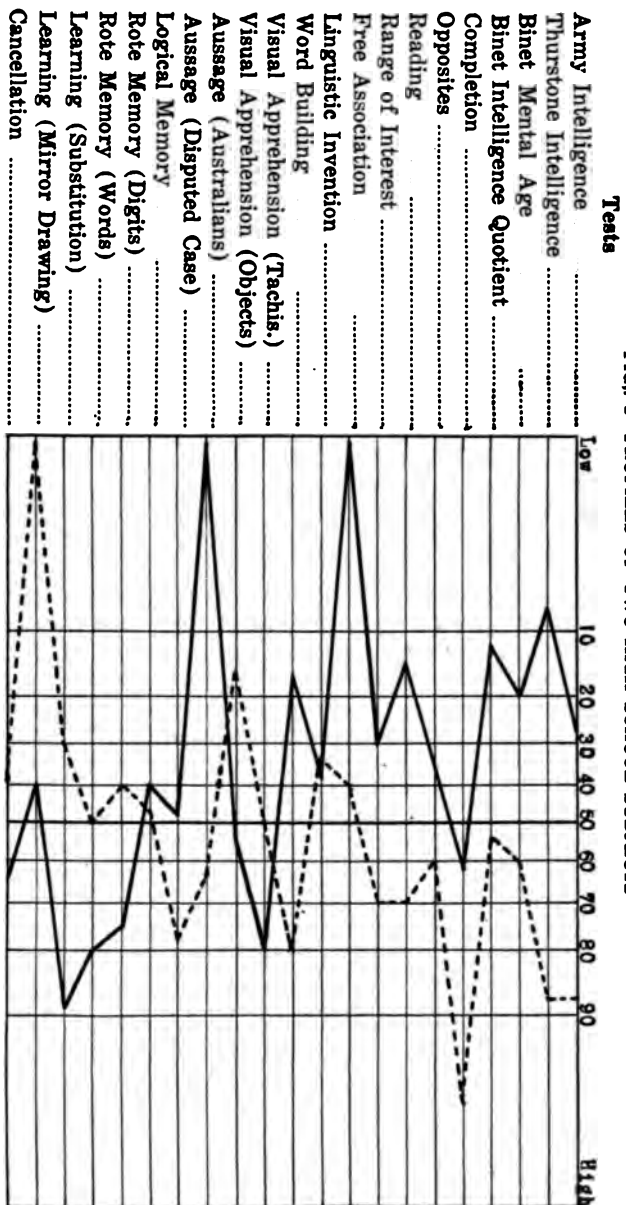
In general ability this girl is approximately at the twentieth percentile. On the average her specific abilities are somewhat better. She makes the lowest record in "Free Association," both written and oral. Her rate of association, which these tests are attempting to measure, is slow. The low record in the "Aussage" Australians is probably accidental, for instead of writing a complete description she merely made a brief list of items. Upon being urged to remedy this defect in the second test she made a record much better than in the first test.

School work of not more than passing grade is expected with such a record. She has earned passing grades in but two commercial subjects and has failed twice in bookkeeping. She realizes her limitations and is taking no more than enough commercial work to secure a position, but will probably fall short of attaining even that limited goal. Her own ill-health and that of her mother keep her out of school a great deal. While she does not have a very positive personality, she is reported to be quite vain about her appearance and spends valuable time in attempting to attract the attention of boys. These efforts are a flat failure. She will soon drop out of school, for she is deriving little benefit therefrom.

Case No. 19. (Variable mental ability.) Boy; age 16-10; freshman; average grade 71; mental age, 17-1; I. Q. 107.

Here is a striking case in which the general ability is average and the specific abilities fall to a much lower level. His highest score is in the "Completion" tests—above the ninetieth percentile. In "Visual Apprehension" he stands below the average, but in the "Aussage" tests, which also involve apprehension, he rises above the average. All his memory test records lie between the fortieth and fiftieth percentiles. His most anomalous record is in the "Mirror-drawing" test, in which his sixth trial record was the same as his first one, the second was even slower than the first, the third and fifth were the highest, and the fourth slightly lower. His efforts throughout this test were expended in trying to reason out what direction to go rather than proceed by trial and error as the instructions dictated. He was so determined on this policy that his record shows decided fluctuations. Being of a mechanical

FIG. 2 PROFILES OF TWO HIGH SCHOOL SUBJECTS



turn of mind he seemed unable to resist the temptation which presented itself to experiment a little.

On the basis of the profile, especially of the general intelligence portion, we would expect him to be a fair student. Nevertheless, he has been one of the most perplexing cases with which the school has had to contend. The boy simply does not study, but spends all his spare time around laboratories, frequently getting numerous minor injuries by unsupervised experimentation. One finger is missing from his left hand, and other injuries are in evidence. He has passed in algebra and General Science, but has repeated his other high-school studies. He takes no interest in the traditional subjects, and spends practically all of his time on work of a mechanical nature. To expostulations about this unwarranted emphasis upon mechanics, he remains unmoved. Attempts to emphasize the fallacy of such an unprecedented waste of time made no impression upon him, although he admitted that he wanted to enter the Engineering College, and described many pieces of apparatus which are found in that college. He spends much time reading "Popular Mechanics," etc., and very little on his school work.

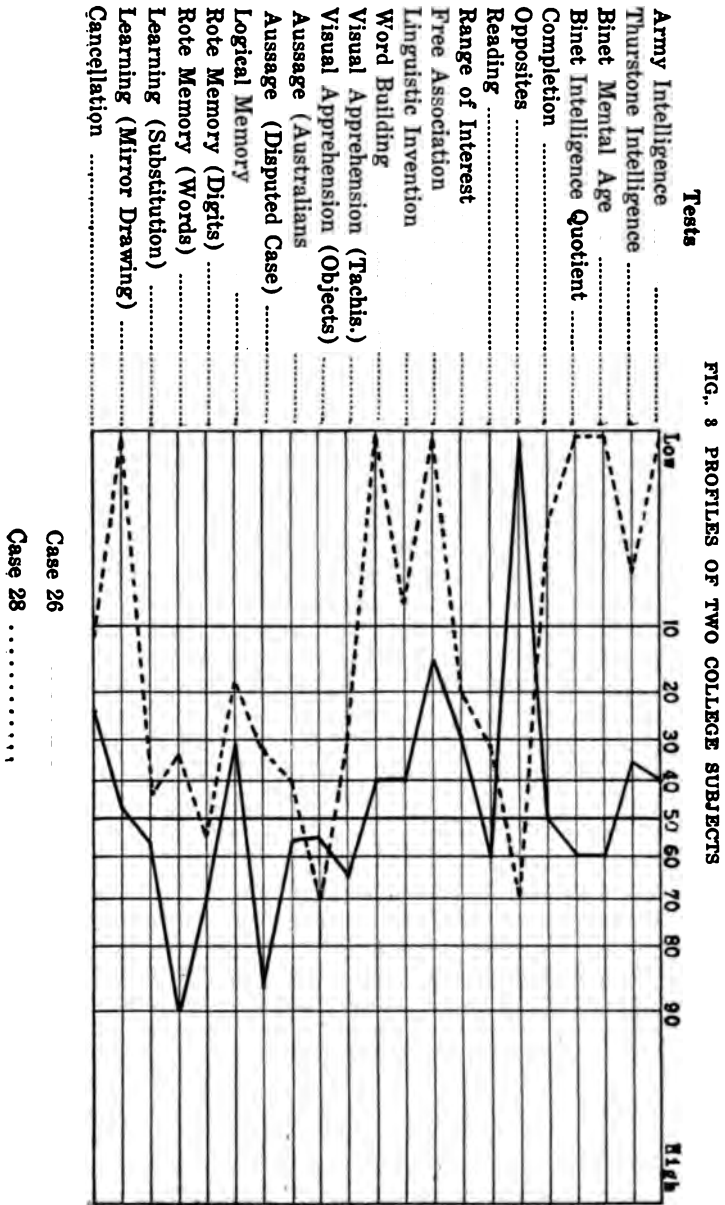
He comes from an inferior family where he is among the older of several children. His eyes are dreamy; his hair is almost white and is wavy. His test profile is strikingly similar to that of case No. 22. It is possibly the result of good ability linked with poor application and little development of specific abilities. In his case unless he can be induced to realize the value of the traditional subjects it would seem well to allow him immediately to work along his natural bent as a major interest. In any event, it is a misfortune to drop such a boy from school, although his record is poor. There should be some way provided by which his ability and interest can be cultivated, for he might easily be successful as a mechanic or inventor. Society needs to conserve its talent in such cases which at least have good chances of success. We regret that our armamentarium of tests did not happen to comprise any that might have measured his mechanical aptitudes.

CASE STUDIES OF THE COLLEGE SUBJECTS

Four college subjects have been selected for discussion. They are No. 26, 28, 43, and 49 of Table IX, in the column headed "Tests."

Case No. 26. (Average and erratic mentality.) Woman; age 18-1; college sophomore; average grade 84 (high D); mental age 18-1; I. Q. 113.

In mentality this woman is well above the average of the general population but not above the average of college students. Her



greatest difficulties are with Associations, as measured by the "Opposites" tests (which are particularly poor), and by "Free Association." She is as low as the thirtieth percentile in "Logical Memory." None of her records is sufficiently superior to place her in the highest tenth of the individuals of this group.

Her college record is also of mediocre quality. She has one B and that is in rhetoric. In several different courses in various fields she has earned C's. In chemistry she has one D and one E.

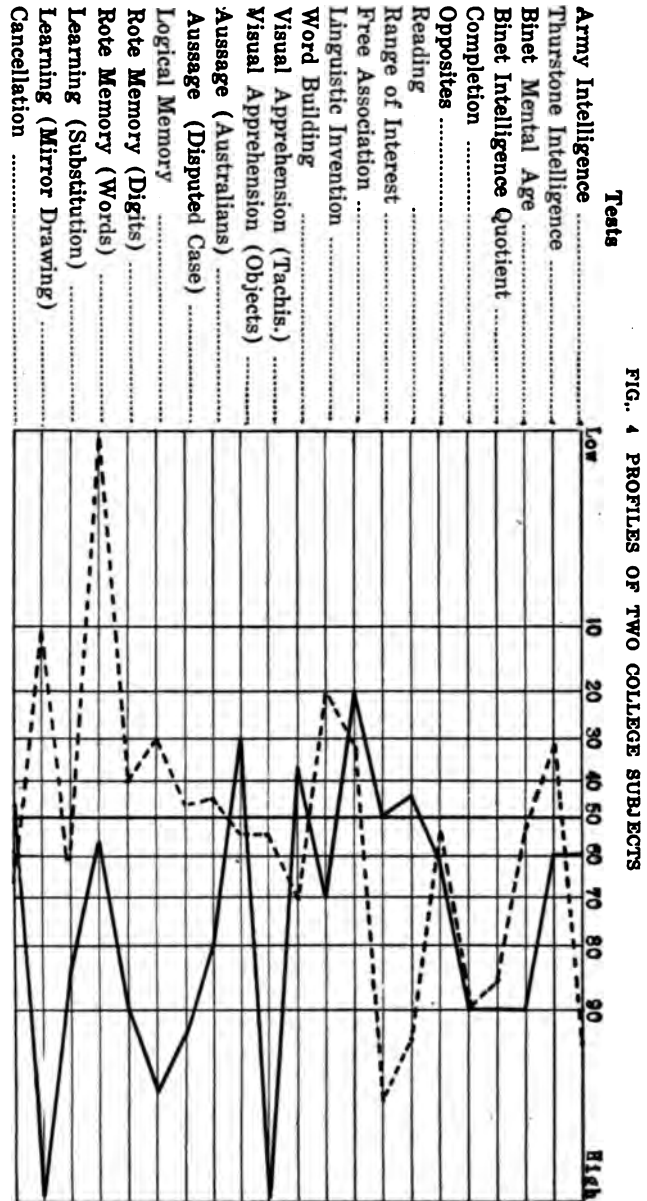
She comes from an average family and has attended the public schools of this city. She has a vague ambition to teach, but has not yet selected her special field of teaching. Even though she should devote herself to study more industriously than is to be expected of a person of this type, she will do well to finish her college course with an average record. The outlook is for generally inferior college work.

Case No. 28. (Consistently inferior mentality.) Man; age 23-4; junior; average grade 66 (high D); mental age 15-10; I. Q. 99.

Poor records and inferior mentality are obviously concomitant in this subject. His score in the "Army" test is only 100 points, our lowest record. In all tests of linguistic ability he is among the lowest five per cent of the college group. His best records are in those tests where quick visual apprehension is an important factor. His memory abilities are next in order of excellence, but are mediocre in comparison with those of other individuals in this group. In "Mirror-drawing" he insisted in trying out the directions in which to move his pencil, although he was frequently warned that success in this test depended upon the method of trial and error. His mental test records throughout are exceedingly poor: he reaches the lowest level of our college group in six tests.

As would be expected, his college records are of the same very inferior quality as his mental ability. He has been retained in college on probation ever since he first entered. His highest grades are C. He has failed signally in several courses in sciences.

He comes from an inferior family. His father is dead and this young man is forced to provide his own means. It is very difficult to be sympathetic with him for his personality is not pleasing; he is extremely strong-headed and opinionated. He expects to enter the Medical college but he will be barred from it by a recent ruling that no student will be accepted in that institution unless he has made an average grade of C in his preliminary college work. Granted that he was admitted, there is no doubt but that he would fail completely. Wise vocational guidance from early youth,



Case 43

Case 49

conducted in the light of facts concerning his mental ability, might well have saved him from his present humiliation, saved him from needless expenditure of time and money, and placed him in a less exacting field where he could have been successful with certain limitations.

Case No. 43. (Consistently superior mentality.) Woman; age 21-10; sophomore; average grade 70 (high D); mental age 19-0; I. Q. 119.

This young woman presents one of the most striking cases which are found in this study. There is a wide discrepancy between her mental ability and her college record. It is evident from her scores in mental tests, as shown on her profile, that she has superior mental ability. Two of her lowest records are in "Visual Apprehension" of objects and in "Cancellation." Her lowest record is in "Free Association," particularly in the oral test. In all but one of the memory tests she has superior records, and she makes excellent scores in both of the "Learning" tests. Her records warrant the belief that she has superior ability.

As a matter of fact, however, her college records are decidedly inferior. Her average grade is lower than C. The crux of her difficulty lies in a curious reticence which amounts to a downright reluctance to answer questions or to talk, at least under classroom conditions. In courses where oral recitation is the most important factor, she has the greatest difficulty. She prepares her lessons quite thoroughly, but is so slow in contributing to classroom discussion that she is passed over as being unprepared, or incapable of mastering the material at hand. Although she is proficient in presenting her ideas in writing, her incompetency in oral recitations more than offsets this proficiency. She now realizes what her chief difficulty is, and in order to improve herself in this respect she is taking oratory and is demanding special attention from her instructor. She declares that this defect did not produce difficulties in her high-school courses.

Although she has had one year's experience in school teaching she has been so completely disheartened at her failure in college that she has not formulated plans for the future, either to enter upon teaching or any other career. She appears to have fallen into a rut from which just now only the special drill in public speaking seems to promise any escape.

Case No. 49. (Average and erratic mentality.) Man; age 22-4; college sophomore; average grade 100 (low C); Mental age 18-1; I. Q. 113.

This man has a better record in tests of general intelligence than in tests of specific abilities. His low record in the "Thurstone

TESTS AND ANALYSIS OF MENTAL CONSTITUTION 371

"Intelligence" test is due to an erroneous idea that there were only 69 problems whereas there were actually 168. He did not hurry in this test until more than half of the time had elapsed, when he suddenly discovered the mistake. In the "Rote Memory" test for words, he became confused where the lists were long, and remembered fewer words than when the lists were somewhat shorter. He became extremely nervous while attempting the "Mirror-drawing" test.

His college record is average. He entered the University this year with advanced standing from a small college where his work had also been of average quality. He stated that he had some difficulty with high-school mathematics, but his "Analogies" test in the "Army" test showed very good performance.

He is an agreeable young man, and comes from a superior family. He intends to specialize in business administration. He was advised to reduce his rate of work and to seek accuracy, in which event he may become moderately successful.

His was a case in which the individual method of testing was exceptionally effective in detecting characteristic peculiarities. He is a rapid worker but uncommonly inaccurate. It was noted, too, that his unwarranted speed of mental operations aggravates his nervous temperament. These conditions were brought to light fully by individual study.

CONCLUSIONS CONCERNING THE MENTAL TESTS

1. Tests specially designed to measure "general intelligence" are, of course, of paramount value in the measurement of all-round mental ability.

2. Tests of specific abilities, however, are essential for anything like an analysis of mental constitution.

3. In order that personal observation might be made, almost all of the mental tests were administered by the individual method, although many of them used in this investigation could well have been administered by the group method, and with a great saving of time. Here might be mentioned, for instance, the "Army Intelligence" tests, and the "Thurstone Intelligence" test.

4. The "Army Intelligence" test although designed for testing a different class of individuals, is nevertheless a very useful test for college and for high-school students. Its value lies chiefly in the splendid selection of material, and in the variety of tests which are included.

5. The "Thurstone Intelligence" test includes items of varying degrees of difficulty, and a penalty might well be imposed upon omissions.

6. The "Stanford-Binet" test is not sufficiently difficult to measure superior adult intelligence. In principle the test is very commendable.

7. The chief criticisms to be leveled at the "Completion" tests used in this investigation is that certain places in them are exceedingly difficult. Other forms now incorporated in various intelligence tests would probably have avoided this difficulty. In general, these tests correlate well with general ability.

8. "Opposites" tests have long been employed and with reliable results. In this study the "Opposites" were too easy for the college subjects, and not only should there have been more difficult terms, but also the number of items might well have been much increased.

9. The "Reading" tests are rightly included in this study, for reading is a common tool of the subject. Poor readers were easily detected by the special complicated text used in reading backwards.

10. The score in the "Range of Interest" test must be interpreted in its relation to maturity in the case of young pupils, but in the case of college subjects it is a direct index to general intelligence with little discount for maturity, save when it needs some discounting on account of unusually wide experience gained in travel or varied occupations.

11. Tests of "Free Association," both written and oral, are very commendable, in that fluency of verbal association is essential to success in them.

12. The "Linguistic Invention" test was of exceptional value in that it presented to all subjects a comparatively novel and interesting task. Ability to compose a short story is a fairly complex ability which taxes productive as well as mere reproductive capacity.

13. "Word-building" tests do not always work as satisfactorily as might be expected, because the more intelligent subjects may waste valuable time in forming useless letter combinations according to some logical plan.

14. The measurement of "Visual Apprehension" by the tachistoscope was generally satisfactory, although the two scores for college students were more consistent than for those for the high-school pupils. Since these tests showed a higher correlation with other tests for the college students than for the high school subjects, we may be led to conclude that they were too difficult for the latter group. Opportunity to handle non-verbal material was the chief value attaching to these tests.

15. Observation, perception and report were measured quite ef-

TESTS AND ANALYSIS OF MENTAL CONSTITUTION 373

fectively by the "Aussage" tests. Although the two test scores were fairly consistent, the slight improvement in the second test was probably due to the practice derived from the first test. These qualities are of no little importance in the student's work.

16. Tests of "Logical Memory" have a slight correlation with general mentality, and are excellent in principle. Tests of "Rote Memory" are, in general, less significant than tests of "Logical Memory" and are useful chiefly for the diagnosis of a rather narrow special ability.

17. "Substitution" is valuable as a learning test if the subjects carry out the intent of the test. The results of this test become more reliable as progress is made from the first to the fifth trial—a point of some interest and one that has not always been observed by users of this type of test.

18. The "Mirror-drawing" test is not only a test of learning by trial and error, but it also is not infrequently an effective means of evoking characteristic temperamental reactions upon the subject's part. Since this test involved a marked emotional element, it is not expected to have a high correlation with other "Learning" tests.

19. The "Cancellation" tests, to our knowledge have never afforded significant correlations with general mental efficiency. Attention might be called to the fact that their highest correlation with other specific tests was in the high-school records, 0.72 with "Substitution" and in the college records 0.32 with the same test. On the other hand, they correlated with the grand total of all mental tests 0.53 in the former and 0.54 in the latter group, respectively. The anomalous fact remains that "Cancellation" shows only low correlation with the "Army Intelligence" test in either group of subjects.

OTHER CONCLUSIONS

20. There is a dearth of reliable norms of mental tests. In order to satisfy requirements fully, percentile tables are most essential. A search through a wide field of literature on mental testing revealed practically no mental test scores which were comparable for our investigation.

21. It is evident from a study of the mental test profiles, and also from the correlation tables that the mentality of a given individual is, on the whole, a uniform affair. That is, high, comparatively average, or low efficiency in any test performance tends to be accompanied by a corresponding grade of efficiency in other performances. This conclusion is obviously quite at variance with the older theory of 'Compensation; and likewise with the prevailing popular notions about the distribution of mental abilities,

though it does accord with the results usually obtained in the laboratory.

22. Many of the seemingly exceptional deviations from this tendency toward a general level ability that appear in certain profiles are to be explained by the operation of disturbing factors, e. g., by a sudden lag of interest, by a misinterpretation of instructions, or by inadequate appreciation of the technique involved.

23. Adaptability is one of the chief assets of superior intelligence. Lack of adaptability is very noticeable among inferior subjects, who were ill at ease in many of the tests involving unfamiliar situations.

24. On all the profiles there is a close correspondence in the scores obtained in the first four tests. The chief exception to this generalization lies in the low records sometimes made by good subjects in the "Thurstone Intelligence" test—records which as already noted, were usually due to a too slow rate of working.

25. Superior intelligence is not an absolute guarantee of excellent scholarship. Lack of interest in subject matter, in some instances lack of industry, or in other instances preference for other activities may contribute to mediocre scholarship.

26. Inferior intelligence is almost universally accompanied by inferior scholarship. By means of untiring industry, inferior intellects may here and there earn average records. But at best their returns are meagre for the energy expended.

27. Average intelligence, we would expect to find accompanied by average scholarship. As a matter of fact, we find it associated with all degrees of scholarship. In these cases, interest and industry become particularly important factors in determining scholarship.

28. Although definite life ambitions were rather clearly revealed in all the subjects examined, it is probably true that the most effective driving motive behind success in school is the ambition to earn a good mark in each course of study pursued.

29. Although home conditions and early training have but negligible influence on mentality, they are essential for the furtherance of ideals and ambitions. In general, better home conditions and better early training tend to be found oftener in pupils of better intelligence. While personality and good character are not necessarily concomitants of superior intelligence, they exert unwarranted influence upon the teachers' estimates of the mental ability of pupils, for instance, a pleasing manner coupled with ready speech and an appearance of really trying, no doubt do frequently mislead teachers with regard to certain inferior pupils.

30. There is a woeful lack of mutual understanding between college students and college instructors. The minor blame for this

TESTS AND ANALYSIS OF MENTAL CONSTITUTION 375

difficulty lies with the instructors, who are unable to cope with personal problems of individual students on account of the large number of students under their direction. The major blame rests upon the students themselves, who seldom confide their troubles to their instructors, doubtless largely for fear that such action will be mistaken for solicitation of special favors in the final record.

31. School work will always run at a less than maximal degree of efficiency until data such as have been presented in this investigation are in the hands of the instructors and full use of those facts is made for each student.

BIBLIOGRAPHY

It has been a difficult task to select a bibliography suitable for this investigation. We are safe in assuming that the standard texts on educational psychology and on the measurement of intelligence should be included. Treatises on the experimental study of mental ability are legion. Many of them deal with the same mental tests which we have just been using. Only a limited number of such references has been included and only those which have some direct bearing upon this study. The chief value derived from such references to periodical literature lies not so much in the norms which are frequently included, as in the suggestions made about the technique of testing.

- (1) J. C. Bell, Mental tests and college freshmen. *Jour. Educ. Psych.* 7: 1916, 381-399.
- (2) M. E. Bickersteth, Mental tests to children of various ages. *Brit. Jour. Psych.*, 9 : 1917, 23-73.
- (3) W. V. Bingham, Some norms of Dartmouth Freshmen. *Jour. Educ. Psych.*, 7 : 1916, 129-142.
- (4) C. Burt, *The Distribution and Relations of Educational Abilities*. London. 1916. Pp. 93.
- (5) L. M. Chassell, Tests for originality. *Jour. Educ. Psych.*, 7 : 1916, 315-323.
- (6) C. B. Cornell, A graduated scale for determining mental age. *Jour. Educ. Psych.*, 8 : 1917, 539-550.
- (7) Edwina A. Cowan, An experiment on the influence of training on memory. *Jour. Educ. Psych.*, 7 : 1916, 31-38.
- (8) R. A. Cummins, Individual differences in a Normal School class. *Psych. Rev.*, 24 : 1917, 403-412.
- (9) A. I. Gates, Correlations of immediate and delayed recall. *Jour. Educ. Psych.*, 9 : 1918, 489-496.
- (10) A. I. Gates, Experiments on the relative efficiency of men and women in memory and reasoning. *Psych. Rev.*, 24 : 1917, 139-146.

- (11) C. Guillet, A study of the memory of young women. *Jour. Educ. Psych.*, 8 : 1917, 65-84.
- (12) F. E. Howard, Psychological differences between children and adults, *Ped. Sem.*, 20 : 1913, 236-253.
- (13) H. L. Hollingsworth, Individual differences before, during, and after practice. *Psych. Rev.*, 21 : 1914, 1-8.
- (14) H. L. Hollingsworth, Variations in efficiency during the working day. *Psych. Rev.*, 21 : 1914, 473-491.
- (15) G. E. Jones, Mental differences between the sexes. *Ped. Sem.*, 20 : 1913, 401-404.
- (16) A. M. Jordan, Some results and correlations of Army Alpha tests. *Sch. and Soc.*, 9 : 1920, 354-358.
- (17) H. D. Kitson, The scientific study of the college student. *Psy. Mon. Suppl.*, 19 : 1917. Pp. 81.
- (18) I. King and T. B. Homan, Logical memory and school grades. *Jour. Educ. Psych.*, 9 : 1918, 266-269.
- (19) Edith F. Mulhall, Tests of the memories of school children. *Jour. Educ. Psych.*, 8 : 1917, 294-302.
- (20) W. H. Pyle, A Manual for the Mental and Physical Examination of School Children. Columbia, 1916. Pp. 32.
- (21) W. H. Pyle, The Examination of School Children. New York, 1913. Pp. 70.
- (22) Henrietta V. Race, A study of a class of children of superior intelligence. *Jour. Educ. Psych.*, 9 : 1918, 91-98.
- (23) G. Rossolimo, Die psychologischen Profile. Zur Methodik der quantitativen Untersuchung der psychischen Vorgänge in normalen und pathologischen Fällen. *Klinik f. psychische u. nervöse Krankheiten*, 6, 1911, Heft 3.
- (24) H. O. Rugg, Statistical Methods Applied to Education. Boston. 1917. Pp. 410.
- (25) D. Starch, Educational Psychology, New York, 1919. Pp. 473.
- (26) D. Starch and I. P. Ash, The mental work curve. *Psych. Rev.*, 24 : 1917, 391-402.
- (27) D. Sunne, The relation of class standing to college tests. *Jour. Educ. Psych.*, 8 : 1917, 193-211.
- (28) L. M. Terman, The Intelligence of School Children. Boston, 1917. Pp. 410.
- (29) L. M. Terman, The Measurement of Intelligence. Boston, 1916. Pp. 362.
- (30) L. M. Terman, The vocabulary test as a measure of intelligence. *Jour. Educ. Psych.*, 9 : 1918, 452-466.
- (31) E. L. Thorndike, Educational Psychology, Briefer Course. New York, 1916. Pp. 442.

TESTS AND ANALYSIS OF MENTAL CONSTITUTION 377

(32) G. R. Wells, An apparatus for the Mirror-drawing test. *Jour. Educ. Psych.*, 9 : 1918, 99-101.

(33) G. M. Whipple, *A Manual of Mental and Physical Tests*. Baltimore, 1914. Part I. Pp. 865; 1915. Part II. Pp. 336.

(34) G. M. Whipple, *Classes for Gifted Children*. Bloomington, 1919. Pp. 151.

(35) Helen T. Wooley, A new scale of mental and physical measurements for adolescents and some of its uses. *Jour. Educ. Psych.*, 6 : 1915, 521-550.

(36) C. S. Yoakum and R. M. Yerkes, *Army Mental Tests*. New York, 1920. Pp. 303.

THE RELATION OF INTELLIGENCE TO AGE IN NEGRO CHILDREN.¹

BY ADA HART ARLITT
Bryn Mawr College

The Army tests have shown negroes of draft age to be inferior to adult whites and a number of investigations, notably that of Strong,² have shown negro children to be inferior to white children. As yet, however, we have no data in the nature of a median I. Q., for all ages will show the extent to which the inferiority of negroes is dependent on age.

Our investigation was undertaken with this point in mind, that is the examination of a sufficient number of children between the ages of five and nine to give at least a tentative norm for each of these ages and of a sufficient number between the ages of ten to fifteen to give a tentative norm of these ages combined.

The subjects were 180 negro children³ tested by the writer in New Orleans during June and July 1921, and 63 tested in the schools in the neighborhood of Philadelphia during the two preceding years. Of the 63 negroes tested in Pennsylvania 36 were examined by the writer and 27 by four graduate students.⁴ Replies in the latter cases were recorded as far as possible verbatim and were graded in consultation with the writer.

The negroes tested in New Orleans were selected in the following manner: 112 came from the negro playground operated by the city, 26 from the Thomy Lafon School, a negro public school taught by negro teachers, 23 from the demonstration classes in a summer normal school operated by Straight University, 10 from a negro private school and 9 from a colored Industrial Home and School. In all cases both as to playground children above 7 years and children tested in the New Orleans schools the children

¹Read before the American Psychological Association, December, 1921.

²Strong, A. C.s Three hundred and fifty white and colored children tested by the Binet Scale. *Ped. Sem.*, 1913, 20: 485-515.

³These subjects were secured through the cooperation of Superintendent Gwinn of the New Public Schools and Mrs. A. J. Stallings of the New Orleans Playground Commission.

⁴For assistance in giving these twenty-seven examinations the writer is indebted to C. Baechele, H. Hanna, I. Neterer and K. Smith.

INTELLIGENCE AND AGE OF NEGRO CHILDREN 379

had been for at least a year previous to the time at which the examinations were made under the instruction of negro teachers. The Pennsylvania negroes were in schools which contained both whites and negroes and had white teachers. The procedure followed in New Orleans was to test every child between the ages of five and nine except in the case of the Thomy Lafon School where the 26 children tested were taken in the order in which they were seated. As they were seated without reference to relative ability this seemed as fair as any other method of random selection.⁵ In the Pennsylvania schools the procedure was to test all negroes between the ages of five and fifteen.

Of the 243 children 129 were boys and 114 were girls. There were 54 of ages five and six combined, 28 boys and 26 girls; 50 seven year olds, 25 boys and 25 girls; 50 eight year olds, 25 boys and 25 girls; 49 nine year olds, 25 boys and 24 girls, and 40 children between the ages of ten to fifteen, 26 boys and 14 girls. As to social status 33.7% were of very inferior, 58% were of inferior, 6.5% were of average, and 1.6% of very superior social status. None were of superior social status.⁶ With four exceptions, two in each of the younger groups, all of the children of average and very superior social status were in the eight, nine and ten to fifteen year old groups.

All were tested by the Stanford Revision of the Binet tests. The school children were tested in a vacant room or hall, the Playground children in a room on the Playground. The conditions for testing were unusually good since all of the children were anxious to take the tests and showed none of the initial shyness so frequently met with in young children. Whenever possible the children were tested in order of age beginning with the youngest. This rule was adhered to strictly in the schools but several exceptions were made in the case of Playground children. We found no evidence of coaching of younger children by older though the few cases of younger children who were tested towards the end of the investigation were closely questioned.

⁵The writer wishes to call attention to an error in results which may easily arise when teachers are allowed to select the children. It is natural that they will select the best, and not the average or poor children, to be tested first. Especially is this the case if there is any question of collecting data for a comparison of two races and the teachers belong to one of the races represented. Unless all the children in the school are tested subsequently, the norm established will represent the superior instead of the average child. In two of the schools tested the teacher was allowed to select the first four children, all eight belonged in the top 25 per cent of the group tested.

⁶The division into social status groups by the occupation of the father was made on the basis of Taussig's five non-competing groups, i. e. unskilled, semi-professional and higher business and professional classes.

The results of the test are to be found in Table I.

TABLE I

I.Q.	Below 60	60-69.9	70-79.9	80-89.9	90-99.9	100-109.9	110-119.9	120-Total	Median	75	25
									Per-	Per-	Per-
									cen-	cen-	cen-
									tile	tile	tile
Age											
5 and 6 yrs.	0	0	1	14	27	9	3	54	100.0	106.6	89.5
7 yrs.	0	3	6	9	30	2	0	50	90.9	96.6	85.3
8 yrs.	0	2	9	15	23	1	0	50	87.5	95.7	81.4
9 yrs.	1	5	8	20	14	0	1	49	83.9	92.2	79.2
10-15 yrs.	1	9	11	15	4	0	0	40	78.4	83.4	71
Total group	2	19	35	73	98	12	4	243	85.6	96.6	81
Percent	.82	7.8	14.4	30	40.3	4.9	1.64				

As the largest percentage of these children came from the inferior social status group our results are most nearly comparable with Terman's results for white children of inferior social status. The median I. Q. for this group of Terman's, all ages combined, is 93. If we may fairly assume this intelligence quotient to be constant for all ages, the I. Q. for negro' children ages five to six combined, which is 100, is 7 points above that of whites of the same social status. Seven year olds have an I. Q. of 90.9 or 2.1 points below that of whites. The median I. Q. for eight year olds is 87.5 or 5.5 points below that of whites, and that of nine year olds is 83.9 or 9.1 points below that of whites, and that of ten to fifteen year olds is 78.9 or 14.1 points below that of whites of the same social status. There is also a corresponding decrease of the 25 and 75 percentiles. At ages five and six combined, the 75 and 25 percentiles are respectively 106.6 and 89.5, at age seven 96.6 and 85.3, at age eight 95.7 and 81.4, at age nine 92.2 and 79.2 and at ages ten to fifteen combined 83.4 and 71.

Combining age groups so that each contains at least 99 individuals we have the following results which are shown in Table II.

The median I.Q. for 104 negro children between the ages of five and seven is 93.5 or .5 of a point above that of native born whites of the same social status, that for 99 eight and nine year olds is 86 or 7 points below that of whites, and that for 139 eight to fifteen year olds is 83.3 or 9.7 points below that of whites of the same social status.

Though we have evidence of a decrease with increasing age of the quotients marking the limits of the middle 50%^a in native born

^aNo data could be obtained as to the proportion of white blood at any age. This factor may play a part in the distribution of intelligence in any of these groups. It is, however, a manifest absurdity to assume that the proportion of white blood is increasingly lower in each age group from the five and six year olds to the ten to fifteen year olds and that this explains the decrease in the median I. Q. and in the 25 and 75 percentiles with increasing age.

TABLE II •

I.Q.	Below 60	60-69.9	70-79.9	80-89.9	90-109.9	110-119.9	120-	Tot.	Med.	75	25
										Per-	Per-
Age										cen-	cen-
5, 6 and 7 yrs. 0		3	7	23	57	11	3	104	93.5	101	88
8 and 9 yrs. 1		7	17	35	37	1	1	99	86	94.7	79.6
8, 9 and 10 to 15 yrs 2		16	28	50	41	1	1	139	83.3	92.3	75.6

whites the median I. Q. remains constant until age 13. Obviously then, age has an effect on the median I. Q. in negro children which it does not have in the case of native born whites. Had the median for the total group, 85.6, been taken as a criterion the negroes would have appeared inferior to whites by 7.4 points, the difference in this case being between an I. Q. which would place a child in the average group and one which would place it in the dull normal group. Actually very young negroes five and six years old have a median I. Q. of 100 or average for the average social status group and 7 points above that of whites of the same social status, while ten to fifteen year olds have a median I. Q. of 78.9 or one which would place a child in what is called by some writers the "border-line" group. There are no superior or very superior children beyond the age of ten and on the other hand the percentage of children with I. Q.s below 70 increases steadily except in the case of the eight year olds from 0% at ages five and six to 25% at ages ten to fifteen combined.

Though the decrease in I. Q. with increasing age is undoubtedly due to a genuine race difference⁸ there are several factors which may also have had an effect.⁹ Negro children are particularly poor in language tests of which there are a much larger percentage at the upper end of the scale than at the lower. The percentage of

⁸The results cited by Terman for native born whites (Stanford Revision of the Binet Tests, page 40, Terman and others, Warwicke and York 1917) show a steady decrease in the quotients at both extremes of the limits including the middle 50 per cent of I. Q.s made at each age, from 97-111 at ages five and six combined to 92-108 at ages eleven and twelve combined.

⁹Odum, H. W. *Annals of the American Academy of Social and Political Science*. Nov., 1913, Vols. 48-50 pp 186-208. Results somewhat similar to ours were printed in this article when the Goddard revision of the Binet tests was applied to negro children in Pennsylvania. On page 102 of the article cited he says: "In the seventh year negro children were approximately as good as the white, and then they decreased to the thirteenth year regularly....." Though his results are open to the same criticisms as those of other investigators using the early revision of the Binet tests, they give added proof that our results are not due merely to some process of selection operative in the region from which most of our cases came.

¹⁰These results do not accord with those of Sunne, (Jour. of Applied Psychol., 1917, p. 82) Sunne states that "Negroes have greater facility in control of words," than have whites.

children between the ages of eight and nine, 99 in all, passing the eight year old vocabulary test is only 33. According to the method by which this test was standardized the percentage of children in the eight year old group passing this test should have been 65 and that in the nine year old group still higher. Our language medium would seem to be too complex for the stage of development which the negro has reached." On the other hand the negro children were particularly good in tests involving rote memory and of these we have more at the lower end of the scale than at the upper. Another factor may be responsible for the superiority of negro children of ages five and six combined. Children of inferior mentality may not have had sufficient initiative to find their way to the Playground or may have been kept home from both Playground and school. Since compulsory education is in force in the districts from which both southern and northern negroes came, this would not apply to any group beyond the age of seven. But since the groups tested are relatively small, this or some other selective factor may have been operative in the five and six year old group.

The effect of sex on the median intelligence quotient is shown in Table III. At ages five and six combined girls are superior to boys; the median for girls is 100.5 and that for boys 94.8. At age seven girls and boys have the same median I. Q. At eight the median for boys, 88, is .5 of a point above that for girls, at nine the median for girls, 84.2 is .9 of a point above that for boys and at ages ten to fifteen inclusive the median for boys, 78.4 is .9 of a point above that for girls.

TABLE III.

I.Q.	Below 60	60- 69.9	70- 79.9	80- 89.9	90- 109.9	110- 119.9	120- 129.9	Median	Total Number
Age 5-6									
Girls	0	0	1	4	13	5	3	100.5	26
Boys	0	0	0	10	14	4	0	94.8	28
Age 7									
Girls	0	1	4	4	15	1	0	90.9	25
Boys	0	2	2	5	15	1	0	90.9	25
Age 8									
Girls	0	1	6	6	11	1	0	87.5	25
Boys	0	1	3	9	12	0	0	88	25
Age 9									
Girls	0	2	4	11	6	0	1	84.2	24
Boys	1	3	4	9	8	0	0	83.3	25
Age 10-15									
Girls	0	4	3	6	1	0	0	77.5	14
Boys	1	5	8	9	3	0	0	78.4	26

"The negro children in the south would tend to have any language difficulty due to race increased since the teachers were of their own race.

INTELLIGENCE AND AGE OF NEGRO CHILDREN 383

Table IV contains the median I. Q. and the distribution of I.Q.s when consecutive ages are combined to give fifty or more children in each group. The median I. Q. for five to seven year old girls is 93.6, that for five to seven year old boys is 93.2 or .4 of a point below that of girls. The median I. Q. for eight to nine year old boys is 86.7 or .7 above that of girls. The median I. Q. for eight to fifteen year old boys is 83.3 or .4 below that of girls. The group medians except at ages five and six combined are practically the same, but at these ages girls are markedly superior to boys. In the total group there are many more girls of superior mentality than boys. Four girls and no boys have I. Q.s above 120 and seven girls and five boys have I. Q.s between 110 and 119.9. Again the proportion of boys of very inferior mentality is much larger than that of girls. Of the twenty-one children with I. Q.s below 70 thirteen are boys and eight are girls. That is to say, there are twice as many girls as boys in the superior to very superior group and almost twice as many boys as girls in the very inferior group.

TABLE IV.

I.Q.	Below 60	60-69.9	70-79.9	80-89.9	90-109.9	110-119.9	120-129.9	Med. 75	Med. 75	25	Total
								Per-	Per-		
								cen-	cen-		
								tile	tile		
Age 5-7											
Girls	0	1	5	8	28	6	3	93.6	102.2	89.7	51
Boys	0	2	2	15	29	5	0	93.2	101.1	85.7	53
Age 8-9											
Girls	0	3	10	17	17	1	1	85.4	94.8	77.3	49
Boys	1	4	7	18	20	0	0	86.7	92.6	80.5	50
Age 8-15											
Girls	0	7	13	23	18	1	1	83.7	92.5	75	63
Boys	2	9	15	27	23	0	0	83.3	92	75.6	76

Our findings as to differences in intelligence between boys and girls are somewhat similar to those of previous investigators. Sunne¹² found that negro girls on the whole were superior to boys by both the Yerkes Point Scale and the Binet tests, though at ages eight, nine and eleven girls were inferior to boys by the Binet tests. They were superior at all ages by the Point Scale. Pyle¹³ found negro girls superior to boys, when averages alone are considered. According to our findings negro girls are superior to boys where either the average I. Q. or the median for the total group is taken. Girls are actually superior to boys only at ages five and six combined and at age nine. The sex difference is very slight at all ages except five and six.

¹²*Ibid.*, cit., page 72

¹³The Mind of the Negro Child, *School and Society* 1915, 1:357-360.

Conclusions.

1. The median I. Q. for negro children decreases with increasing age from age six to ages ten to fifteen combined.

2. At ages five and six negroes are superior to whites of the same social status. At all ages beyond six negroes are inferior to whites and this inferiority increases with increasing age.

3. Though there are more girls than boys with intelligence quotients above 110 and more boys than girls with intelligence quotients below 70, the differences due to sex are slight except at ages five and six. At these ages girls are markedly superior to boys.

THE COMPUTATION OF PEARSON'S r FROM RANKED DATA

BY CLARK L. HULL
University of Wisconsin

I.

Psychologists frequently desire a reliable coefficient of correlation where one or both of the arrays of data consist of ranks. This is particularly the case in the evaluation and weighting of vocational tests where often the only obtainable criterion is a series of ranks, though all the test data will usually be in ordinary units. Spearman's rank-difference coefficient of correlation can not be used satisfactorily as a basis for regression equations¹ and even if it were so used, the prediction of a criterion in terms of a more or less accidental series of ranks would have little meaning. A method of obtaining Pearson's r from ranks must accordingly be found if such situations are to yield readily to the vocational psychologist.

The proposed method is, in brief, to convert the series of ranks into an array of scores on a convenient linear scale, the linear scores then to be used in the computation of r in the ordinary way. In translating the ranks into linear units, it is assumed that the ability in question is distributed according to the Gaussian law. It then becomes easy with the aid of a table computed from that of Sheppard² to tell the percent of cases falling between the various successive points of the scale. For example, if the range of ranked ability be assumed arbitrarily as cut off at 2.5σ in each direction, and that the limits of a 10-point scale exactly coincide with the respective extremes of this range, the highest 1.7% of the ranked cases will fall between scores 9 and 10. The next 4.4% of the cases will fall between 8 and 9, the next 9.3% between 7 and 8, the next 15.2% between 6 and 7, the next 19.4% between 5 and 6, and so on. These values being known, it is a simple matter to take the first 1.7% of any given series of ranks and assign to them the central value of the highest scale interval, to assign to the next highest 4.4% of ranks the central value of the next highest scale interval, and so on. Thus if n be 60, 1.7% of this series yields in round numbers, 1 case, which accordingly receives the score of 9.5; 4.4% of 60 yields 3 cases at 8.5; 9.3% yields 5 cases at 7.5. Table I shows the data suggested above, worked out fully and arranged in a systematic manner.

¹Kelley, Truman Lee. *Tables: To Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations*, p. 24.

²Pearson, Kari. *Tables for Statisticians and Biometricians*, pp. 2-8.

TABLE I

Score values on a linear scale	Per cent of cases to be given each score	Number of cases receiving each score where n is 60
(A)	(B)	(C)
9.5	1.7%	1
8.5	4.4%	3
7.5	9.3%	5
6.5	15.2%	9
5.5	19.4%	12
4.5	19.4%	12
3.5	15.2%	9
2.5	9.3%	5
1.5	4.4%	3
.5	1.7%	1
Mean, 5.0	Total, 100.0%	Total, 60

II

It may be desired for some purpose to translate ranked data into scores of a linear scale by individual ranks instead of by groups as above. The method of doing this and the table used, depend upon the same principles as the method just described. If c represent the rank of a given subject in a series and n the number of subjects ranked, then the formula

$$\% = \frac{100 (c - .5)}{n}$$

yields values such as are given in column A of Table II. This figure is practically a percentage expression of the distance in rank of a given subject from the good or efficient extreme of the distribution of ability in question. Column B of this table gives the corresponding score on a 10-point scale by steps of .1 point.

The operation of this second method may be illustrated by a sample computation. Suppose n to be 25 and c to be 1. Substituting in the above equation we have,

$$\begin{aligned} \% &= \frac{100(1 - .5)}{25} \\ &= 2 \end{aligned}$$

COMPUTATION PEARSON'S r FROM RANKED DATA 387

Looking up 2 in column A of Table II, we find the value nearest it to be 1.96. In column B opposite this entry is found the corresponding score or linear scale value, 8.9. If there are no fractional ranks in the series, the % values after rank 1 may be determined by merely adding successively $\frac{100}{n}$ (in the above case, 4) until the whole series of ranks has been traversed.

TABLE II

% (A)	Score (B)	% (A)	Score (B)	% (A)	Score (B)
.09	9.9	22.32	6.5	83.31	3.1
.20	9.8	23.88	6.4	84.56	3.0
.32	9.7	25.48	6.3	85.75	2.9
.45	9.6	27.15	6.2	86.89	2.8
.61	9.5	28.86	6.1	87.96	2.7
.78	9.4	30.61	6.0	88.97	2.6
.97	9.3	32.42	5.9	89.94	2.5
1.18	9.2	34.25	5.8	90.83	2.4
1.42	9.1	36.15	5.7	91.67	2.3
1.68	9.0	38.06	5.6	92.45	2.2
1.96	8.9	40.01	5.5	93.19	2.1
2.28	8.8	41.97	5.4	93.86	2.0
2.63	8.7	43.97	5.3	94.49	1.9
3.01	8.6	45.97	5.2	95.08	1.8
3.43	8.5	47.98	5.1	95.62	1.7
3.89	8.4	50.00	5.0	96.11	1.6
4.38	8.3	52.02	4.9	96.57	1.5
4.92	8.2	54.03	4.8	96.99	1.4
5.51	8.1	56.03	4.7	97.37	1.3
6.14	8.0	58.03	4.6	97.72	1.2
6.81	7.9	59.99	4.5	98.04	1.1
7.55	7.8	61.94	4.4	98.32	1.0
8.33	7.7	63.85	4.3	98.58	.9
9.17	7.6	65.75	4.2	98.82	.8
10.06	7.5	67.48	4.1	99.03	.7
11.03	7.4	69.39	4.0	99.22	.6
12.04	7.3	71.14	3.9	99.39	.5
13.11	7.2	72.85	3.8	99.55	.4
14.25	7.1	74.52	3.7	99.68	.3
15.44	7.0	76.12	3.6	99.80	.2
16.69	6.9	77.68	3.5	99.91	.1
18.01	6.8	79.17	3.4	100.00	.0
19.39	6.7	80.61	3.3		
20.83	6.6	81.99	3.2		

In case the $\frac{100}{n}$ should turn out to involve a large decimal, the additions may be carried out very easily on an ordinary adding or computing machine. The results of the conversion of a series of ranks by the method just described, is shown in Table III.

TABLE III

Rank	%	Score
1	2	8.9
2	6	8.0
3	10	7.5
4	14	7.1
5	18	6.8
6	22	6.5
7	26	6.3
8	30	6.1
9	34	5.8
10	38	5.6
11	42	5.4
12	46	5.2
13	50	5.0
14	54	4.8
15	58	4.6
16	62	4.4
17	66	4.2
18	70	3.9
19	74	3.7
20	78	3.5
21	82	3.2
22	86	2.9
23	90	2.5
24	94	2.0
25	98	1.1

The first column shows a series of 25 ranks. The second column shows these ranks converted into % values by the formula. The third column shows the corresponding score on a 10-point scale to the nearest .1 point, as given in Table II.

III

The above arbitrary choice of a scale of 10 steps into the units of which all ranked criteria are to be converted, calls for some comment. That the scale should be arbitrary is obviously inevitable. A scale of 10 points was chosen because of its simplicity and general intelligibility. If desired, however, the decimal point may be disregarded in the scale values, in which case there results a 100-point scale. A 10-point scale has the additional advantage that each step corresponds to a fairly definite concept of ability. A convenient set of ability concepts is shown in the following table, each concept opposite its corresponding step of the scale.

COMPUTATION PEARSON'S r FROM RANKED DATA 389

Scale unit	Ability concept
10	maximal
9	excellent
8	very good
7	good
6	above average
5	average
4	below average
3	poor
2	very poor
1	inferior
0	minimal

It may even be that in certain cases, particularly in industry, final prediction in terms of ability concepts would be more acceptable than in numerical scale units.

Occasions will sometimes arise where it will be desirable to make predictions in terms of some other scale such, for example, as the ordinary percentage scale of school marks. In such cases the ranks should first be converted into units of the 10-point scale by one of the methods described above. These scores may then be converted into the units of any scale that may be desired by the special conversion formula mentioned below.

Despite the occasional necessity of predicting criteria in terms of the most divers units, in vocational guidance there is a distinct practical advantage in making predictions in terms of some simple uniform scale such as that suggested above. Ideally the vocational adviser should be able to tell the subject which of all the vocations in the world he is best fitted to pursue. Practically he ought to be able to tell the subject *relatively* how much aptitude he has for a limited number of vocations. This demands a separate prediction on each vocation, while comparability of predicted aptitude demands a uniform prediction scale for all. Take a typical vocational group such as typing, stenography, bookkeeping and general office work. Ordinarily the criterion of the first would be the number of words typed per minute, the criterion of the second would be the number of words written down per minute, the bookkeeping might be in terms of school marks, while the original criterion at least, on office workers, might be in terms of ranks in a group of 60, say. Now suppose, with a view to discovering which of the four vocations would be most likely to bring success to a given individual, his aptitude in each were predicted by means of appropriate regression equations. At least to the uninitiated, it is by no means obvious whether a certain number of words typed per minute would be a better or worse score than some other number of words writ-

ten in shorthand, or whether either of these would be better or worse than a certain school grade or a certain rank in a group of 60, to say nothing of how much better or worse. As the result of this lack of natural comparability of the various prediction scales, the vocation of greatest promise might still be far from evident after the labor of making the predictions had been performed. If, on the other hand, prediction on all four criteria had been made in terms of a single uniform scale, the relative promise of the various vocations would be instantly evident to any one and the object of making the predictions would be attained.

There is no technical obstacle in the way of the adoption by psychologists of a uniform criterion scale for purposes of vocational prognosis. The simplicity of the procedure where the criterion originally has been obtained in terms of ranks, has been described. In many cases the criterion will be obtained in terms of some linear scale such as the number of words typed per minute or the average number of boxes made per day in a factory. Such scores may readily be converted into units of any standard uniform scale desired, by means of a special conversion formula. This will be found in Vol. VI., p. 298f of this Journal, together with examples of its use. In using this formula to convert linear scores into units of a 10-point scale, the mean of the 10-point series should be taken as 5 and the standard deviation as 2. These values will need to be substituted in the conversion formula.

A COMPARISON OF THREE TESTS OF "GENERAL INTELLIGENCE"

BY MORRIS S. VITELES,
University of Pennsylvania

This study concerns itself with the examination of fifty-nine students of the Wharton School of Finance and Commerce of the University of Pennsylvania with three tests of "general intelligence." The purpose of the study is, 1: to measure the variability in performance of the members of a selected, homogenous group in a number of accepted tests of general intelligence, 2: to measure the relationship between success in each of these tests and school grades, the measures of academic success.

METHOD.

The Otis General Intelligence Test, the Army Alpha and the Morgan Mental Test were used in this investigation. The tests were given within a period of two months in the order in which they are mentioned. They were all given between eleven and twelve in the morning to the members of a class assembled for a lecture in Psychology. The men taking the tests were juniors and seniors in the Wharton School of Finance and Commerce of the University of Pennsylvania.

The average of the grades in all courses for a period of the academic year preceding the one in which the tests were given was taken as the measure of academic success. The grades given in the University of Pennsylvania are D (Distinguished), G (Good), P (Passed), N (Not passed) and F (Failed). A student who receives N in a course is entitled to a re-examination; a student who receives F is not allowed to take a re-examination but must repeat the course. The following numerical values were assigned to each of these grades:

D equals 85
G equals 75
P equals 65
N equals 55
F equals 35

For each student the total value of the grades for the year was obtained by multiplying each grade by the value assigned to it and finding the sum. This total was divided by the number of courses to give the average grade for the year. The minimum number

of courses taken during the year by any student was found to be 8: the maximum number was found to be 20 and the average 15.5.

RESULTS.

A TREATMENT AND ANALYSIS

1.—The variability of the performance in the test was first investigated by a study of the rank order displacements of the individuals in the group when one test is compared with the others. The results of this analysis are presented in Table 1, below:

TABLE I
TABLE OF RANK DIFFERENCES
INTELLIGENCE TESTS.

	Alpha-Otis		Alpha-Morgan		Otis-Morgan	
	No.	%	No.	%	No.	%
0- 5.5	30	50.9	20	33.9	17	28.9
6-10.5	10	16.9	11	18.7	15	25.5
11-15.5	8	13.6	12	20.3	7	11.8
16-20.5	6	10.0	6	10.0	6	10.0
21-25.5	2	3.4	5	8.5	6	10.0
26-30.5	2	3.4	2	3.4	4	6.8
31-35.5	1	1.8	2	3.4	1	1.8
36-40.5	0	0	0	0	2	3.4
41-45.5	0	0	0	0	1	1.8
46-50.5	0	0	1	1.8	0	0
51-55.5	0	0	0	0	0	0
56-58-	0	0	0	0	0	0
Total	59	100.0	59	100.0	59	100.0

The maximum displacement in rank of any given individual on the Army Alpha and Otis Test is 32.0. The minimum is 1.0. The average displacement is 8.1 and the median displacement is 5.5. The correlation of the tests by the rank difference method (Pearson formula) gives a coefficient of plus 0.76.

Between the results of the same individual on the Alpha Test and the Morgan Test there is greater discrepancy, on the average, than between the Army and the Otis Tests. The maximum rank order displacement between these two tests is 49.5; the minimum is 0.0. The average displacement is 11.2 and the median is 10.0. The co-efficient of rank and order correlation is plus 0.62.

The greatest rank order displacement exists between the Morgan Test and the Otis Test. The maximum displacement in this case is 46.0; the minimum is 0.0. The average is 12.3 and the median is 8.0. The value of the rank order co-efficient of correlation between these two tests is plus 0.51.

This analysis shows that there is less variability in the performance of the individuals of a selected group with the Army Alpha and the Otis Test than between Alpha and the Morgan Test. The greatest variability in performance comes when the same individ-

COMPARISON OF THREE INTELLIGENCE SCALES 393

uals are given the Morgan Test and the Otis Test, the analysis of individual displacements and the low value of the co-efficient of rank and order correlation indicating that the chance for a given individual to be assigned to closely related positions in rank order distribution of the results of the two tests is a small one.

2.—The second method of measuring the extent of variability in the performance of a group in a series of tests is *a* to divide the group into tertiles, higher, middle and lower and to compute the number of individuals assigned to the same tertile by the different tests; and *b* to divide the group into quartiles and compute in the same way, the number and percentage of individuals assigned to the same quartiles by the different tests. The results of such an analysis are presented in Charts 1 and 2.

CHART I COMPARISON OF TERTILE ASSIGNMENTS

Alpha-Otis-Morgan Tests.

	Alpha - Otis	Alpha - Morgan	Morgan - Otis																																																
Alpha	<table border="1" style="border-collapse: collapse; margin: auto;"> <tr><td>I</td><td>1</td><td>5</td><td>14</td></tr> <tr><td>II</td><td>5</td><td>10</td><td>4</td></tr> <tr><td>III</td><td>13</td><td>7</td><td>0</td></tr> <tr><td></td><td>III</td><td>II</td><td>I</td></tr> </table>	I	1	5	14	II	5	10	4	III	13	7	0		III	II	I	<table border="1" style="border-collapse: collapse; margin: auto;"> <tr><td>I</td><td>0</td><td>8</td><td>12</td></tr> <tr><td>II</td><td>9</td><td>7</td><td>4</td></tr> <tr><td>III</td><td>11</td><td>6</td><td>2</td></tr> <tr><td></td><td>III</td><td>II</td><td>I</td></tr> </table>	I	0	8	12	II	9	7	4	III	11	6	2		III	II	I	<table border="1" style="border-collapse: collapse; margin: auto;"> <tr><td>I</td><td>1</td><td>5</td><td>12</td></tr> <tr><td>II</td><td>10</td><td>10</td><td>2</td></tr> <tr><td>III</td><td>9</td><td>6</td><td>4</td></tr> <tr><td></td><td>III</td><td>II</td><td>I</td></tr> </table>	I	1	5	12	II	10	10	2	III	9	6	4		III	II	I
I	1	5	14																																																
II	5	10	4																																																
III	13	7	0																																																
	III	II	I																																																
I	0	8	12																																																
II	9	7	4																																																
III	11	6	2																																																
	III	II	I																																																
I	1	5	12																																																
II	10	10	2																																																
III	9	6	4																																																
	III	II	I																																																
	Otis	Morgan	Morgan																																																
	62.7 per cent. in same tertile	50.8 per cent in same tertile	52.5 per cent. in same tertile																																																

The chance that a given individual will be assigned to the same tertile by any two tests is $33\frac{1}{3}$ percent. The highest percentage of individuals assigned to the same tertile by any two tests is 62.7% (Alpha and Otis), only twice the percentage assigned by chance. The lowest assignment is 52.8%, or only one and a half times the chance assignment. (Alpha and Morgan.)

Chart 1 shows not only the percentage of individuals assigned to the same tertile by two different tests but the detailed distribution among the tertiles. The first figure in Chart 1 shows that out of the twenty individuals assigned to 1, or the lower tertile in the Alpha Test, 1 or 5% is found in the 3 or highest tertile in the Otis Test; 5 or 25% are found in the middle tertile according to the Otis results and the remaining 14, or 70% are found in the 1, or lower tertile in the distribution of the results of the examination with the Otis Test. On the other hand by reading vertically up

from the figures on the abscissa it can be found that out of the 13 individuals assigned to 3 or the higher tertile by the Otis Test 13 or 68% are found in 3 or the higher tertile in the distribution of the Alpha results; 5 or 26.8% are found in the middle tertile in the distribution of the Alpha results and 1 or 5.2% in 1 or the lower tertile.

The remaining figures in this chart, and in the charts similar to it presented elsewhere in this report show in the same way the number of individuals of a given tertile or quartile in the distribution of results on one measure assigned to the various tertiles or quartiles in the distribution of the results on another measure.

b/—In addition to the division of the group into tertiles on each of the tests a division was made into quartiles. The results of this analysis are presented in Chart 2.

CHART II

COMPARISON OF QUARTILE ASSIGNMENTS

Alpha-Otis-Morgan Tests.

Alpha-Otis		Alpha-Morgan		Morgan-Otis	
Alpha	I	0	0	5	8
	II	1	6	5	4
	III	2	6	4	2
	IV	12	3	1	0
	IV	III	II	I	
Otis		Morgan		Morgan	
52.5 per cent.		40.6 per cent.		44.0 per cent.	
in same quartile		in same quartile		in same quartile	

The chance that any individual will be assigned to the same quartile by any two tests is 25%. The greatest percentage of individuals assigned to the same quartile by any two tests is 52.5% (Alpha and Otis). The smallest percentage is 40.6% (Alpha and Morgan). The chart shows the number of individuals assigned to a given quartile in a given test assigned to each of the quartiles by the other tests.

A survey of this analysis shows that with reference to the variability in the assignment to tertile and quartile the greatest difference exists between the Alpha and the Morgan Tests and the least difference between the Alpha and the Otis. Otis and Morgan show almost as much variability in the assignment of given individuals to tertile and quartile as do Alpha and Morgan. Only

COMPARISON OF THREE INTELLIGENCE SCALES 395

in the case of the tertile assignment in Alpha and Otis is the percentage of assignments to same tertiles high enough to show that the tests measure somewhat similar abilities.

3.—The variability in the performance of a group of individuals a number of tests may be due to a disagreement among the authors of the test on the question of what constitutes general intelligence. As a result of such disagreements different material, testing different abilities, and weighted differently, may be included in the test. The extent to which the tests agree in the measurement of the same intelligence will be indicated in the co-efficient of correlation. Such co-efficients of correlation have been computed for these tests. These co-efficients, computed by the product moment method, are presented in Table 2, and Table 3.

TABLE II

Coefficients of Correlations Alpha, Morgan and Otis Tests	
Morgan-Alpha	$r\ 0.52 \pm .071$
Morgan-Otis	$r\ 0.59 \pm .061$
Alpha-Otis	$r\ 0.69 \pm .048$

TABLE III

Partial Co-efficients of Correlation	
If Morgan equals 1	
If Alpha equals 2	
If Otis equals 3	
$r_{12.3}$	plus $0.21 \pm .091$
$r_{13.2}$	plus $0.37 \pm .080$
$r_{23.1}$	plus $0.54 \pm .07$

That similar elements are present in all three tests is revealed both by the value of the co-efficients of correlation between the tests and by the fact that when the influence of the elements of one test present in the other two is removed in the correlation of the other two tests, by the method of partial correlation, the resultant co-efficient of correlation between the two tests is greatly diminished. There is, however, a difference in the degree of similarity between the elements of the three tests. Alpha and Otis are most similar in the elements contained, with a co-efficient of correlation high enough to be of considerable significance. Morgan and Alpha are least similar. As a matter of fact the elimination of the elements common to Otis in the correlation of Morgan and Alpha reduces the co-efficient of correlation of Morgan and Alpha to plus .21±.091, much below the level of significance. In the same way the elimination of the elements common to Alpha in the correlation of Morgan and Otis reduces the co-efficient of

correlation of these two tests to plus $0.37 \pm .080$, also below the level of significance. The elimination of the Morgan elements in the correlation of Otis and Alpha does not affect seriously the value of the co-efficient of correlation between these two tests, demonstrating more conclusively the fact that, to a great similar extent elements are tested by Otis and Alpha. It is by reason of the similarity of the elements tested by the two that there is least variability in the performance of the group in these two tests.

B.—COMPARISON OF TEST RESULTS WITH SCHOOL GRADES.

The second part of the study was designed to measure the relationship between success in each of these tests and success in school, as represented in school grades. This part of the study was naturally suggested by the increased use of tests of "general intelligence" as criteria for college entrance, and purposes to give some slight additional evidence, of different accepted measures of "general intelligence" in predicting academic success.

The statistical treatment of these results parallels that given above, but in addition to test results, school grades are introduced and become the objective standard or criterion of comparison in each case.

1.—The difference in rank order assignment of the individuals of the group in each test as compared with school grades is given in Table IV below.

TABLE IV

TABLE OF RANK DIFFERENCES, INTELLIGENCE TESTS AND SCHOOL GRADES.

Units of Rank Difference	Alpha No.	School %	Otis No.	School %	Morgan No.	School %
0- 5.5	13	22.0	12	20.3	13	22.0
6-10.5	7	11.8	8	13.5	7	11.8
11-15.5	10	16.9	10	16.9	12	20.3
16-20.5	9	15.2	6	10.0	4	6.8
21-25.5	5	8.5	8	13.6	8	13.6
26-30.5	5	8.5	8	13.6	3	5.1
31-35.5	3	5.1	2	3.4	5	8.5
36-40.5	2	3.4	3	5.1	0	0
41-45.5	4	6.8	1	1.8	3	5.1
46-50.5	1	1.8	1	1.8	4	6.8
51-55.5						
Total	59	100.0	59	100.0	59	100.00

The maximum displacement of any individual in rank order when the Otis scores are compared with school grades is 49.0, the minimum being 0.5. The average displacement of the group is 17.4; the medium is 15.5. The co-efficient of correlation, by the rank order method, is plus 0.22.

Between the Alpha Test and the School grades the maximum displacement is 48.0. The minimum is 0.5. The average is 17.4

COMPARISON OF THREE INTELLIGENCE SCALES 397

and the median is 15.5. The co-efficient of rank order correlation is plus 0.22.

In the comparison of Morgan Scores with school grades even greater discrepancy in rank order placement is revealed. The maximum displacement in this case is 49.0 and the minimum 0.0. The average is 20.8 and the median 15.5. The co-efficient of correlation by the rank order method between the two criteria is plus 0.13.

These results indicate that there is no greatly significant relationship between the scores on any of the tests and school grades. The average displacement in rank order between each of the test scores and the school grades is high, and the co-efficient of correlation is in each case, low. Of the three tests the Otis scores show the greatest relationship with school grades and the Morgan scores the least.

2. In order further to examine the relationship between success in each of these tests with academic success an analysis was made of the variability in tertile and quartile assignments of each test as compared with tertile and quartile assignments made on the basis of an array of school grades.

a/—The analysis of tertile assignments is presented in Chart 3 below.

CHART III

COMPARISON OF TERTILE ASSIGNMENTS

Tests and School Grades

Morgan - School Grades					Alpha - School Grades					Otis - School Grades				
Morgan	I	7	5	6	Alpha	I	5	7	8	Otis	I	4	5	9
	II	6	8	7		II	7	6	6		II	8	7	7
	III	8	6	6		III	9	6	5		III	9	7	3
		III	II	I			III	II	I			III	II	I
School Grade					School Grade					School Grade				
in same tertile 37.3 per cent.					in same tertile 38.9 per cent.					in same tertile 42.3 per cent				

In view of the well defined tendency of school administrators to divide children of the same chronological age or of the same peda-

gological age into three classes, on a basis of differences in mental endowment, for purposes of education, the comparison of tertile displacements between test scores and school grades is particularly significant. On the basis of chance alone 33⅓% of those examined would be assigned to the same tertile by both school grades and test scores. In the charts above, as it has been stated before, the numbers enclosed in heavy lines represent the number of individuals assigned to the same tertile both by school grades and test scores.

By the Morgan test and school grades 37.3% of the individuals examined are assigned to the same tertile; by the Army Alpha Test and school grades 38.9%; by Otis scores and school grades 42.3%. In each case the variability in assignment to tertiles is high, the percentage of parallel assignments being in no case very much higher than the chance assignment. Of the three, however, the Otis scores parallel most closely the assignment by school grades.

b/—In Chart 4 below is presented the analysis of quartile assignments when tests scores are compared with school grades.

CHART IV
COMPARISON OF QUARTILE ASSIGNMENTS
Tests and School Grades

Morgan - School Grades		Alpha - School Grades		Otis - School Grades	
Morgan	I	3	3	2	6
	II	5	4	5	1
	III	4	5	4	3
	IV	5	4	3	4
	IV	III	II	I	
School Grades		School Grades		School Grades	
32.2 per cent.		32.2 per cent		28.8 per cent.	
in same quartile		in same quartile		in same quartile	

There is a twenty-five percent chance that individuals will be assigned to the same quartile by a given test score and school grade. In the case of the Morgan test 32.2% of those tested are placed in the same quartile to which they are assigned in the distribution of school grades. The same percentage, 32.2%, are assigned to the same quartiles by Alpha score and school grades, and 28.8% by Otis scores and school grade.

The charts also show how many individuals assigned to a given

tertile or quartile by test is assigned to each of the school grade quartiles and conversely, the number of those assigned to a given tertile or quartile by school grade assigned to each of the tertiles in the distribution of test results.

A summary of the results shows that no well defined tendency for those rating either high or low in school grades to rate high or low in any of the tests. There is a vague suggestion of such a grouping in the case of the Otis Test and school grades, 50% of those in Tertile 1, the lower tertile in this test being in the lower tertile in school grades, and 47% of those in the higher tertile in the test being in the higher tertile in school grades, but even here the percentages are not particularly significant, and they are less significant in the case of the other tests.

3.—In order to determine whether similar elements are involved in success in each of these tests and academic success in the Wharton School as portrayed in school grades, the scores in each test were correlated with school grades by the product moment method. The co-efficient of correlation so obtained are presented in Tables 5 and 6 below.

TABLE V.

Co-efficients of Correlation	
Test Scores and School Grades	
Morgan-School Grades	r plus 0.06
Alpha-School Grades	r plus 0.21
Otis-School Grades	r plus 0.24

TABLE VI.

Partial Co-efficients	
If Morgan equals 1	
If Alpha equals 2	
If Otis equals 3	
If School G equals 4	
r 14:23 equals	—0 .09
r 24:13 equals	+0. 04
r 34:12 equals	+0 .11
By the method of multiple correlation	
R 4:123 equals	Plus 0.252

In no case is the correlation between test scores and school grades significant. The elimination of the elements common to the other two tests in the correlation of any test with school grades diminishes the value of the co-efficient of correlation of the test with the school grades. Altho all the co-efficients are insig-

nificant in value the highest co-efficient is given by the correlation of Otis scores with school grades.

The correlation of the combined tests with school grades by the method of multiple correlation gives a co-efficient of plus 0.252, indicating that not even the battery of three tests has great significance in the prediction of academic success in the Wharton School.

CONCLUSIONS.

1.—The discrepancy in the results of individual members of the group on the different tests of "general intelligence" leads to the conclusion that the mental ability measured by each of these tests is not the same. They are all called tests of "general intelligence" but it cannot be concluded from this investigation that there is a similar complex or pattern of abilities measured by the three tests.

2.—Greater similarity in the individual elements of the test exists between the Otis and Alpha tests than between the Morgan tests and either of the first two. With regard to the mental factors involved in performance, the Morgan test and the Alpha test show the greatest dissimilarity.

3.—The comparison of test results with academic grades indicates that none of these three tests of "general intelligence" can be used in the prediction of academic standing in the Wharton School. It is possible, of course, that in the examination of an unselected group of applicants for entrance into the Wharton School the potential failures might be picked by one or all of these tests, but the statistical analysis presented in this report throws no light upon this problem.

DISCUSSION OF RESULTS.

The most important problem raised by these results is that of determining which of these three tests measures that which we are calling "general intelligence." It is quite possible that this problem will not be solved until we are able to agree upon a definition of "general intelligence." No attempt will be made to give such a definition in this article or to determine which of these tests is the best measure of "general intelligence." The purpose of this article is merely to point out the great discrepancy in the standing of the individual members of a group on three tests which are presumably designed to measure the same mental ability. The fact that there is greater discrepancy in the individual standing of the

members of the group when the Morgan test is compared with the Alpha and Otis than when these two tests are compared with each other, is in itself, no indication that the Morgan test is less a test of "general intelligence" than are the other two tests. The fact that none of the tests correlate with academic standing is not necessarily a sign that any or all of the tests do or do not measure "general intelligence." The finding of this study are only of importance in so far as they point out the need for evaluating very carefully tests which are proposed for the measurement of "general intelligence" for purposes of education, selection of workers in industry, or for general experimental purposes.

The absence of significant correlation, rank order assignment, and tertile and quartile assignment between test results and school grades is significant only in so far as it points out the necessity for determining exactly to what extent the abilities required for success in any test of "general intelligence" are factors in determining academic standing in any of the schools of a University. It is quite possible that the ability measured in the tests used in this investigation are not the same as those required for success in a specialized school of business such as the Wharton School of Commerce and Finance of the University of Pennsylvania. A careful comparison between test results and the success and failure of individuals selected for each of the specialized schools of the University is necessary in order to determine the value of a given test of "general intelligence" as part of the Entrance Requirements of any University. The comparison of test results and academic standing of the pooled data from the different schools of the University is not sufficient. The pattern or group of abilities or the general competency required in one school is undoubtedly different than those required in another, and a test of "general intelligence" which is applied successfully in the prediction of academic success in one school may be altogether useless for prediction of academic success in another. It may be necessary for the psychologist to introduce in his thinking on this subject a concept of what might be called **specialized general intelligence** or **specialized performance pattern** or **ability-group-pattern** representing the pattern of specific mental ability necessary for success in a given specialized field such as engineering, business, etc., —the professions such as distinguished from the high commercial jobs, etc. That there is some recognition of such a problem is indicated in the appearance of such tests as the "Thurstone Test for Engineering Aptitude" and the "Otis Test of General Intelligence for Business Institutions" and the reported plan of the

nificant in value the highest co-efficient is given by the correlation of Otis scores with school grades.

The correlation of the combined tests with school grades by the method of multiple correlation gives a co-efficient of plus 0.252, indicating that not even the battery of three tests has great significance in the prediction of academic success in the Wharton School.

CONCLUSIONS.

1.—The discrepancy in the results of individual members of the group on the different tests of "general intelligence" leads to the conclusion that the mental ability measured by each of these tests is not the same. They are all called tests of "general intelligence" but it cannot be concluded from this investigation that there is a similar complex or pattern of abilities measured by the three tests.

2.—Greater similarity in the individual elements of the test exists between the Otis and Alpha tests than between the Morgan tests and either of the first two. With regard to the mental factors involved in performance, the Morgan test and the Alpha test show the greatest dissimilarity.

3.—The comparison of test results with academic grades indicates that none of these three tests of "general intelligence" can be used in the prediction of academic standing in the Wharton School. It is possible, of course, that in the examination of an unselected group of applicants for entrance into the Wharton School the potential failures might be picked by one or all of these tests, but the statistical analysis presented in this report throws no light upon this problem.

DISCUSSION OF RESULTS.

The most important problem raised by these results is that of determining which of these three tests measures that which we are calling "general intelligence." It is quite possible that this problem will not be solved until we are able to agree upon a definition of "general intelligence." No attempt will be made to give such a definition in this article or to determine which of these tests is the best measure of "general intelligence." The purpose of this article is merely to point out the great discrepancy in the standing of the individual members of a group on three tests which are presumably designed to measure the same mental ability. The fact that there is greater discrepancy in the individual standing of the

members of the group when the Morgan test is compared with the Alpha and Otis than when these two tests are compared with each other, is in itself, no indication that the Morgan test is less a test of "general intelligence" than are the other two tests. The fact that none of the tests correlate with academic standing is not necessarily a sign that any or all of the tests do or do not measure "general intelligence." The findings of this study are only of importance in so far as they point out the need for evaluating very carefully tests which are proposed for the measurement of "general intelligence" for purposes of education, selection of workers in industry, or for general experimental purposes.

The absence of significant correlation, rank order assignment, and tertile and quartile assignment between test results and school grades is significant only in so far as it points out the necessity for determining exactly to what extent the abilities required for success in any test of "general intelligence" are factors in determining academic standing in any of the schools of a University. It is quite possible that the ability measured in the tests used in this investigation are not the same as those required for success in a specialized school of business such as the Wharton School of Commerce and Finance of the University of Pennsylvania. A careful comparison between test results and the success and failure of individuals selected for each of the specialized schools of the University is necessary in order to determine the value of a given test of "general intelligence" as part of the Entrance Requirements of any University. The comparison of test results and academic standing of the pooled data from the different schools of the University is not sufficient. The pattern or group of abilities or the general competency required in one school is undoubtedly different than those required in another, and a test of "general intelligence" which is applied successfully in the prediction of academic success in one school may be altogether useless for prediction of academic success in another. It may be necessary for the psychologist to introduce in his thinking on this subject a concept of what might be called *specialized general intelligence* or *specialized performance pattern* or *ability-group-pattern* representing the pattern of specific mental ability necessary for success in a given specialized field such as engineering, business, etc., —the professions such as distinguished from the high commercial jobs, etc. That there is some recognition of such a problem is indicated in the appearance of such tests as the "Thurstone Test for Engineering Aptitude" and the "Otis Test of General Intelligence for Business Institutions" and the reported plan of the

School of Business Administration of Harvard to use a special test of "general intelligence" as part of its entrance requirements.

The desirability of making further investigations along this line is pointed out to some extent by the absence of correlation between the test results and academic standing in the Wharton School. In view of the purpose of this article merely to present some additional statistics on three tests of "general intelligence" in order to point out the need of evaluating more carefully proposed measures of "general intelligence" the further discussion of this problem will be left to another article.

GROUP INTELLIGENCE EXAMINATIONS FOR PRIMARY PUPILS

O. J. JOHNSON

Division of Research, Public Schools, St. Paul, Minnesota.

Part I.

A BRIEF SURVEY OF EXISTING EXAMINATIONS.

In 1919 the first successful attempt was made to construct group intelligence examinations for primary pupils. During this year four primer tests were published. One of these, the preliminary form of the present Haggerty Intelligence Examination Delta 1^{*} was given to a large number of pupils in the Virginia School Survey. Almost contemporaneous with it, came the publication of the Lowell Group Intelligence Scale, the Pressey Primer Scale, and the Myers Mental Measure. Since this time the activity in this field has been great and a number of other examinations have appeared at odd intervals. At the present time, there are about fifteen examinations designed to test the mental abilities of pupils from kindergarten through the fourth grade. These tests, although all intended for group use, vary widely in the kinds of exercises they contain, the total number of items, the length of time taken to give, directions, methods of scoring, and accessory materials. They differ even more in the apparent care exercised in construction, the completeness of statistical data, and useableness as practical instruments in the hands of the teacher.

Some of these examinations are not on the market, nor is there any published literature describing them. The list given below is believed to contain all that have been devised, and they have been arranged in the order of their appearance so far as this has been possible to ascertain.¹

(1919) Haggerty Intelligence Examination, Delta 1^{*}

(1919) Lowell Group Intelligence Scale for Primary Grades²

(1919) Myer's Mental Measure³

¹ Haggerty, M. E., *Virginia Public Schools*, Part 2, pp. 115-119, 1921.

² A new group test by Herschel T. Manuel has appeared since the study was completed.

³ Haggerty, M. E., *Hagerty Intelligence Examination*, World Book Co.

⁴ Lowell, Frances, *A Group Scale of Intelligence for Primary Grades*, J. App. Psych., Vol. 3 pp. 215-247, 1919.

⁵ Myers, Caroline E. and Garry C., *A Group Intelligence Test*, Sch. and Soc., Vol. 10, pp. 355-360, 1919.

- (1919) Pressey Primer Scale⁶
 Towne Picture-Game Test⁷
 Detroit First Grade Test⁸
- (1920) Holley Picture Completion Test⁹
- (1920) Otis Group Intelligence Scale, Primary Examination Forms A and B¹⁰
 Seattle Entrance Class Tests¹¹
- (1920) Smith Graded Intelligence Tests¹²
- (1920) Dearborn Group Tests of Intelligence, Series 1¹³
- (1921) Johnson Non-Verbal 2¹⁴
- (1921) Cole-Vincent Group Intelligence Test for School Entrance¹⁵
- (1921) Kingsbury Primary Group Intelligence Scale, Form A¹⁶

It has been very difficult to gather information of value about several of the examinations mentioned in this study. This was especially true in the case of statistical data. Because of this fact, the comparisons made in Table I are confined to material gathered largely from test booklets and from directions for giving and scoring. In this table the first column shows how widely different are the periods of time taken to give the tests, the extremes being 8 and 85 minutes. The next two columns contrast the actual material of which the examinations are composed. For this, the number of items is a better measure than the number of tests, as the latter differ enormously in length. The table also shows the tremendous influence of Binet in the choice of test material. The lack of fore-exercises in several instances is to be deplored, whereas the high objectivity in scoring is very encouraging, both from the standpoint of efficiency and reliability.

⁶ Pressy, Luella Winifred, *A Group Scale of Intelligence for Use in the First Three Grades*, J. Ed. Psych., Vol. 10, pp. 297-308, 1919; and J. Ed. Research, Vol. 1, pp. 285-294, 1920.

⁷ Towne, Clara Harrison, *Towne Picture-Game Test*, Published by the Orthogenic Clinic, Public Schools, Kansas City, Mo.

⁸ Engel, Anna M., *Detroit First-Grade Intelligence Test*, World Book Co.

⁹ Holley, Charles E., *A Picture Completion Test*, described in *Mental Tests for School Use*, Univ. of Ill. Bulletin No. 28, Vol. 17, March 8, 1920.

¹⁰ Otis, Arthur S., *Otis Group Intelligence Scale, Primary Examination*, World Book Co.

¹¹ *Seattle Entrance Class Tests*, Board of Education, Seattle, Wash.

¹² Smith, Franklin, O., *Graded Intelligence Tests*, University of Montana.

¹³ Dearborn, W. F., *Dearborn Group Tests of Intelligence, Series 1*, J. B. Lippincott and Co.

¹⁴ Johnson, O. J., *Non-Verbal 2 Intelligence Examination for Primary Pupils*, Dept. of Education, Public Schools, St. Paul, Minn.

¹⁵ Cole L. W., and Vincent, Leona E., *Group Intelligence Tests for School Entrance*, State Normal School, Emporia, Kansas.

¹⁶ Kingsbury, Forest A., *Kingsbury Primary Group Intelligence Scale*, Public School Publishing Co.

TABLE I
Giving Comparisons Between Primary Intelligence Examinations

Name of Examination	No. Min. to Give	No. Tests	Total No. of Items	No. New Tests Not in Other Exams	Tests Adapted from Binet	No. Tests Having Foreign Exercises	For What Grades Exam is Adapted	No. Tests Objectively Scored	No. Tests Requiring Charts To Give
Lowell	70	25	84	11	11	0	1-2-3	25	2
Myers Mental Measure	20	3	51	1	1	3	1-2-3	3	1
Pressey Primer Scale	25	4	100	0	1	4	1-2-3	4	0
Seattle	48	4	67	3	0	4	1 (low)	4	2
Towne	23	13	26	0	11	6	1 (low)	13	0
Haggerty Delta 1	30	6	119	2	3	6	1-2-3	5	0
Detroit	24	15	64	0	4	0	1	14	0
Holley	8	1	20	0	1	1	1-2-3	1	0
Otis	45	8	100	3	4	8	1-2-3	8	0
Smith	85	10	100	2	2	10	1-2-3-4	10	0
Dearborn	85	14	197	10	10	0	1-2-3	12	5
Non-Verbal 2	40	5	94	1	3	5	1-2-3	5	1
Cole-Vincent	40	8	46	1	2	4	1 (low)	6	2
Kingsbury	22	4	46	2	1	4	1-2-3	4	0
Average	37.5	7.8	79.5						

TABLE II
SHOWING NUMBER OF TIMES EACH KIND OF TEST
APPEARS IN 14 PRIMARY EXAMINATIONS.

Names of Tests Fre- quencies		Names of Tests Fre- quencies	
1. Oral Directions	11	27. Mixed-up Sentences	1
2. Picture Completion	10	28. Word Classification	1
3. Substitution	7	29. Moral Judgment	1
4. Classification of Pictures	6	30. Identifying Eyes, Nose and Mouth	1
5. Drawing Designs	5	31. Recognition of Sex	1
6. Pictorial Identities	5	32. Telling Time	1
7. Tapping	4	33. Body Attitude	1
8. Length of Lines	4	34. Naming Pictures	1
9. Word Comparison	3	35. "Dominoes"—Numerical Quantities	1
10. Maze	3	36. Dictation	1
11. Flashboard	3	37. Cossing out Letters	1
12. Absurdities	3	38. Perception of Sound	1
13. Geometrical Figures	3	39. Number of Toes	1
14. Counting Objects	3	40. Tying Bow Knot	1
15. Dot Pattern	2	41. Writing from Dictation	1
16. Spot Card	2	42. Comprehension	1
17. Color Naming	2	43. Drinking Cup	1
18. Asthetic Comparison	2	44. Tapping Squares	1
19. Color Form	2	45. Alphabet Test	1
20. Ball and Field	2	46. Memory of Words	1
21. Value of Stamps	2	47. Opposites	1
22. Picture Sequence	2	48. Completion	1
23. Pattern Completion	2		
24. Pictorial Relations	2		
25. Memory for Digits	2		
26. Common Sense	1		

DISTRIBUTION OF TYPES OF TESTS IN PRIMARY EXAMINATIONS.

In Table II all tests have been classified into groups. One cannot fail being impressed by the large number of different kinds of exercises that have been devised to measure the intelligence of young children. It is interesting to note, also, the popularity of Oral Directions and Picture Completions. Of these tests, twelve

GROUP INTELLIGENCE TESTS FOR PRIMARY GRADES 407

are directly adapted from the Binet-Simon Scale, or its revisions. Eight others are derived from the same source, but have been greatly modified. The remaining twenty-eight are drawn from group intelligence examinations for older persons, or have been invented as new tests.

CORRELATIONS WITH BINET

Table III gives the results, so far as it has been possible to ascertain them, of the correlations between various primary examinations and Binet mental ages. The differences in composition of the groups of pupils used in determining these relations make direct comparison of the coefficients a very difficult matter, but they are at least suggestive.

TABLE III
SHOWING CORRELATION OF PRIMARY EXAMINATIONS
WITH BINET MENTAL AGES.¹⁷

Examination	r	No. and Kind of Cases
Haggerty Delta ¹⁸	.74	27 (Public school)
Lowell	.76	76 (Public school)
Myers Mental Measure	.81 to .83	Less than 30 cases (Public school)
Pressey	.75	64 (Feeble-minded children)
Pressey	.66	148 (Public school)
Seattle ¹⁹	.738	274 (Entering school)
Seattle	.447	97 (Entering school)
Dearborn ²⁰	.73	140 (Feeble-minded)
Dearborn	.81	88 (Public school)
Non-Verbal 2	.765	63 (Subnormal pupils)

Part II.

The Non-Verbal 2 Intelligence Examination for Primary Pupils.²¹

SELECTION AND CONSTRUCTION OF TESTS.

The first work on the Examination was begun by the writer early in 1919. It consisted of a careful study of all non-verbal tests with a view of selecting such as could be readily adapted to young children. After discarding a number which proved undesir-

¹⁷ These correlations are found in the articles previously referred to.

¹⁸ Statistical computation done by Eleanor Bremer, Mental Hygiene Association, Philadelphia, Pa.

¹⁹ Letter to writer Jan. 30, 1921.

²⁰ Letter to writer Feb. 11, 1921.

²¹ This Examination is called Non-Verbal 2 to distinguish it from Non-Verbal 1 for kindergarten and first grade now in preparation.

able, the following were retained: Part Wholes, Maze, Number and Letter Checking, Geometrical Forms, Oral Directions, Designs, Substitution, Pictorial Completion, and Pictorial Identities. After repeated trials and revisions, it was necessary to discard the first four of these tests—the Part Wholes because it was verbal²² and the others because of the difficulty of adapting them to first grade pupils. The last five were again revised and put into as usable form as possible. These are the tests of which the Non-Verbal 2 Examination is composed.

The method of selecting or rejecting items was as follows. All tests were given to about seventy first and second grade children. The papers were then arranged in the order of scores on all the tests combined. Each item was subjected to the following treatment: if more pupils standing above the median in all tests passed an item successfully than pupils standing below, the item was considered satisfactory in respect to significance, and if acceptable in other ways, was retained. In this manner, it was possible to eliminate items not discriminative of the kind of abilities measured by the whole examination. In addition to this, each item was also checked up against Stanford-Binet scores of sixty-five first grade pupils. In the case of a few doubtful items, the combined scores of all tests and Stanford-Binet mental ages were used for final disposition.

When all undesirable items had been thrown out, the remaining ones were arranged in order of difficulty, using the P.E. as a measure²³. In a few instances where large gaps were found, other items of the desired difficulty were inserted in the required places, thus insuring a more even gradation from the easiest to the most difficult.

In order to give a more comprehensive idea of the kind of tests included in the Non-Verbal 2 Intelligence Examination, a brief description of each will be given below. This will be supplemented by comparisons with other tests in order to point out the main differences or similarities.

THE ORAL DIRECTIONS TEST.

This test differs from other tests in use mainly in regard to length. It contains sixteen items graded in difficulty and four others which are used as a fore-exercise. Great length is of importance in that it tends to increase reliability, thus making the scores of much greater value. For this reason and because an exercise of this sort serves to adapt pupils to the whole testing

²² This test was used as a part of the criterion described farther on.

²³ The tables from which the P.E. values were calculated are given in Woody, Clifford, *Measurements of Some Achievements in Arithmetic*, Teachers Coll. Contrib. to Educ., No. 80, p. 87.

GROUP INTELLIGENCE TESTS FOR PRIMARY GRADES 409

situation and particularly to the much longer uninterrupted working periods of the tests that follow, it was given first place in the examination.

Considered from one standpoint, directions tests fall into two distinct classes. Some items call for scarcely any specific knowledge of facts, while others verge on being tests of general information. A test like Abelson's "Geometric Figures" ²⁴ is of the former kind and calls mainly for correct performance of the directions, while in the following example, there is needed specific knowledge of a particular kind: "If a captain is superior to a corporal, put a cross in the second circle." Here the stress is laid on the information.

In making up the material to go into the Oral Directions Test of the Non-Verbal 2 Examination care was exercised that it should belong to neither of these two types. Some information of a practical nature is called for, but in most cases correct responses depend on the understanding, and ability to execute, the commands, with speed.

THE DESIGNS TEST.

Designs tests also naturally divide themselves into two groups. In the first, the subject copies the design while looking at it; in the second, it is reproduced immediately after being observed. Both kinds of tests were used by Binet and Simon and have later been included in almost all revisions of their scale. But the majority of such exercises have one conspicuous disadvantage, namely, that the designs drawn by the children take multifarious shapes which are very hard to score. To remedy this defect, the figures used in Non-Verbal 2 have been so constructed as to be distinctively objective in scoring. By this plan the pupil must remember the design, but his response takes definite shape, because part of the figure is already on the paper. His work is to fill in the missing part. Such drawings can also be evaluated objectively by counting the missing lines which the pupil has supplied.

The Substitution Test.

It is likely that in the case of young children the most important abilities measured by this test are those involved in making the substitutions with speed rather than in the rapidity of learning the symbols. The time limit of three minutes is too short to permit any except the brightest pupils to learn the symbols, but experimentation seems to indicate that it is extremely difficult to secure maximum effort from primary pupils for a period longer than this. Most tests now used are far briefer and hence fall short of secur-

²⁴ Abelson, A. R., *Mental Ability of Backward Children*, Br. Jr. of Psych., Vol. 4, pp. 286-314, 1911.

ing an adequate indication of a child's ability. A unique feature of the exercises in the Non-Verbal 2 is the simplicity of the symbols. All of them, except one, can be made with one stroke of the pencil. Most symbols in other examinations require two or even three strokes and this unduly retards young children who have not yet developed adequate control over hand and arm muscles. By making the signs easy, it was believed that these defects could be minimized thus enhancing the value of the exercise. Its high reliability and significance as shown by statistical analysis bears out this hypothesis.

THE PICTURE COMPLETION TEST.

Picture completions as found in tests may be divided into two groups. In the most common kind, a part of an object is missing. Here the matter is factual and successful performance depends to a considerable extent on acquaintance with the object represented. This becomes, then, virtually a test of general information. In the second kind, the picture represents a natural principle; i. e., two animals, one of which casts a shadow while the other one does not. The third kind of item is a symmetrical design, a part which has been left out.

The picture completion test in the Non-Verbal 2 Examination belongs to the first and third types. All items are either symmetrical designs, or pictures of objects with which children are, in general, familiar. It may be stated that the symmetrical design is a form of exercise, original with the author, which shows promise of usefulness.²⁵ The designs included in this test have proven excellent; they are easy to give and objective to score.

THE PICTORIAL IDENTITIES TEST.

This form of exercise is one of the latest to develop, and it appears to fill a unique place.²⁶ When one examines group tests for young children, one is struck with the fact that none appear to call for the higher types of thinking. They deal more with sensory abilities, perception, recognition of things, motor manipulation, and so on, rather than the more conceptual processes involved in abstract thinking. The predominance of the first mentioned kind of tests for primary pupils is undoubtedly proper, inasmuch as young children have not yet developed the higher thought activities and hence should not be given many tests measuring them. But these processes are not entirely absent even in early childhood and they seem to become more predominant

²⁵ This kind of item appears also in the Otis Primary Examination.

²⁶ This test was made up by selecting twenty items from the preliminary forms of Non-Verbal B published by the National Research Council. The Pictorial Identities Tests in these examinations were originally devised by Dr. M. E. Haggerty to be used in the Virginia School Survey. They are included in the Non-Verbal 2 Examination by his permission.

GROUP INTELLIGENCE TESTS FOR PRIMARY GRADES 411

as children grow older. They also appear to be fundamental for human progress." For these reasons, at least one such test ought to be included in an intelligence examination, and it seems that Pictorial Identities fills the place.

CRITERION OF INTELLIGENCE.

In order to determine how well an examination measures the traits it is supposed to measure, it is necessary to evaluate it in terms of some adequate criterion. This important procedure in test construction seems to have been neglected by many workers. It is a thing that cannot be determined offhand, but must be ascertained through rather laborious statistical methods. The early work in establishing criteria made almost exclusive use of ratings of pupils by their teachers and fellow students. For the final criterion, the separate judgments of several persons were averaged to form a more reliable measure. In recent years greater weight has been given to more objective data, such as composite scores on several kinds of tests, grade location, age, and scholarship of pupils as shown by teachers' marks.²⁷ By these methods more comprehensive evaluations of pupils' native mental capacities result, and they tend to be more valuable because they give less weight to non-intellectual traits.

The criterion against which the significance of the scores obtained by the use of the Non-Verbal 2 Examination was evaluated is given below. It consists of the pooled results of the following five measures of pupils' mental abilities: the scores on two non-verbal intelligence examinations, two reading tests, and teachers ratings of intelligence. The ratings were given numerical values. All of these measures were then added together to form a composite whole and the correlations between it and the Non-Verbal 2 Examination were determined for each separate test and the Total Score. The criterion is given below:

CRITERION.

1. Haggerty Intelligence Examination, Delta 1.
2. Pressey Primer Scale (Scores times one-half.)
3. Haggerty Reading Examination, Sigma 1.
4. Part Wholes Test (Vocabulary)²⁸
5. Teachers' Ratings, using these designations to which were assigned the accompanying number values: Excellent (16); Very Good (12); Good (8); Fair (4); Poor (0).

²⁷ Terman, L., (32) *Intelligence and Its Measurement: A Symposium*, J. Ed. Psy., Vol. 12, pp. 127-133, 1921. (Article by L. M. Terman)

²⁸ See Haggerty, M. E., *Haggerty Intelligence Examination, Manual of Directions*, and Terman, L. M., *Terman Group Test of Mental Ability, Manual of Directions*, World Book Company.

²⁹ Mentioned at the beginning of this article. The test contains words selected from first and second grade readers in common use.

TABLE IV

Showing the Correlation of the Separate Tests and of the Total Score of the Non-Verbal 2 Examination with Criterion.

(101 pupils, grades 1 high to 3 high).

Tests	r	P.E.
Total Score (Present five Tests)	.841	.019
Oral Directions	.673	.020
Designs	.546	.046
Substitution	.739	.030
Pictorial Completion	.503	.049
Pictorial Identities	.636	.039
Geometrical Figures	.278	.061
Total Score, Six Tests	.822	.020

In Table V below the significance of the various combinations of tests has been calculated in a way believed to be original with the writer. The plan was to see if the coefficients of correlation with the Criterion would increase or decrease when tests were omitted from the examination. The results show that the coefficient decreases when any one of the first five tests are omitted, but that it increases when the Geometrical Figures test is left out. This indicates that all the tests but this one add elements which help to increase the correlation of the whole examination with the Criterion. As a result of this drop in the coefficient, the Geometrical Figures Test was left out.

TABLE V

Showing the Correlation with Criterion of the Six Tests First Included in the Non-Verbal 2 Examination. It Further Shows the Change in the Size of the Coefficient Due to Leaving out Successively the scores of one of these tests. (101 pupils, grades 1 high to 3 high.)

Tests	r	P.E.
Total Score, Six Tests	.822	.020
These Six Tests, but		
1. Leaving out Oral Directions	.795	.024
2. Leaving out Designs	.807	.023
3. Leaving out Substitution	.777	.026
4. Leaving out Pict. Compl.	.810	.022
5. Leaving out Pict. Ident.	.797	.024
6. Leaving out Geom. Figures	.841	.019

GROUP INTELLIGENCE TESTS FOR PRIMARY GRADES 413

CORRELATION WITH STANFORD-BINET

Opportunity was found to give the tests to a group of 63 subnormal pupils who had previously been Bineted. These children ranged in mental age from 5 years 9 months to 13 years and had been examined during a period stretching back somewhat over two years. The correlation coefficients are given in Table VI. They are not as high as one might at first expect. Due to the difficulty in giving group tests to subnormal pupils and because of the unreliability of the mental ages themselves, and the lapse of time, in many cases, since the individual tests were given, a high degree of correspondence between the scores can not be expected.

TABLE VI

Showing Correlation of Non-Verbal 2 Examination with Mental Ages (63 subnormal pupils).

Test	r	P.E.
Total Score (Present five tests).	.765	.035
Oral Directions	.625	.051
Designs	.473	.065
Substitutions	.743	.037
Pictorial Completion	.356	.073
Pictorial Identities	.610	.052

TABLE VII

Showing Correlation of Separate Tests of Non-Verbal 2 Examination with Mental Ages on the Stanford-Binet Scale (107 Cases, Grades 1 high to 3 high).

	r	P.E.
Oral Directions	.598	.040
Designs	.519	.046
Substitution	.706	.032
Pictorial Completion	.630	.038
Pictorial Identities	.757	.023

INTERCORRELATIONS.

The statement made in the preceding section that tests should correlate closely with the Total Score needs one qualification. It seems desirable only when intercorrelations between the tests themselves are low. If an exercise correlates high with the other tests, its close correlation with the Total Score would be due to its measuring the same abilities as the other tests. Low correlations on the other hand, indicate that the exercises, although meas-

uring essentially different functions, unite on the one point of testing the same traits as the examination itself.

Table VII reveals that the tests in Non-Verbal 2 measure essentially different mental functions. The last column gives the average of the intercorrelation coefficient for each test and indicates how each is related to the other four. In no case is the relation close.

TABLE VIII

Showing Intercorrelations of Tests of Non-Verbal 2 Examination (107 Cases, Grade 1 high to 3 high).

Tests	Oral Direct.	De-signs	Substi-tution	Pict. Comp.	Pict. Ident.	Aver-
Oral Directions r		.195	.361	.272	.372	.300
P.E.		.063	.056	.059	.055	
Designs r	.195		.216	.225	.251	.219
P.E.	.062		.061	.060	.059	
Substitution r	.361	.216		.380	.357	.328
P.E.	.056	.061		.054	.056	
Pictorial r	.272	.225	.380		.467	.336
Completion P.E.	.059	.060	.054		.050	
Pictorial r	.372	.241	.357	.467		.359
Identities P.E.	.055	.060	.056	.050		

RELIABILITY OF TESTS.

The definition given by Burt is so comprehensive that it will be repeated here.³⁰ He says: "The amount of fluctuation to which a given mode of measurement is liable can be determined by finding the correlation between two sets of experimental figures obtained on different occasions or by different observers for one and the same measurement, or the average correlation if the measurement was applied more than twice. The coefficient thus obtained is called the Reliability Coefficient, and measures inversely the reduction of apparent correlation caused by errors of observation incidental to that measurement." This method has been used by several investigators, and there can be no doubt that it is of the greatest importance to every one who tries to devise new tests of any sort, even though it has certain limitations as recently pointed out by Kelley.³¹

In the case of the Non-Verbal 2, the tests were given to the pupils with an interval of one week. The medians of each test for the entire group were calculated for the two givings to deter-

³⁰ Burt, Cyril, *Experimental Tests of General Intelligence*, Br. Jr. Psych., Vol. 3, pp. 95-177, 1909.

³¹ Kelley, Truman Lee, *Reliability of Tests*, J. Ed. Research, Vol. III, pp. 370-379, May 1921.

GROUP INTELLIGENCE TESTS FOR PRIMARY GRADES 415

mine the amount of improvement. It was remarkable that out of the 107 pupils only one made a lower score the second time.

TABLE IX

Showing Median Improvement in Tests of NonVerbal 2 Examination. (107 Cases, Grades 1 high to 3 high)

	Median Score First Giving	Median Score Second Giving	Improvement in per cent
Oral Directions	9	11	22
Designs	9	12	33
Substitutions	8	11	37
Pictorial Comp.	10	11	31
Pictorial Ident.	9	12	33
Total Score	44	54	24

It is contrary to popular opinion that so very constant results can be obtained in the examination of young children by the group method. The coefficients in Table X prove that with proper care the group method is almost as adequate for primary grades as with the upper grades or in high school.

TABLE X

Giving Reliability Coefficients of Non-Verbal 2 Examination. (107 Cases, Grades 1 high to 3 high)

	r	P.E.
Total Score	.832	.019
Oral Directions	.560	.044
Designs	.794	.023
Substitution	.707	.032
Pictorial Completion	.730	.030
Pictorial Identities	.802	.022

CORRELATION WITH SCHOLARSHIP.

Scores in Non-Verbal 2, and of three other Intelligence examinations were correlated with teacher's ratings of scholarship. These ratings were made by persons who had taught the children one semester or more. The scheme used was a five group classification using the designations: Excellent, Very Good, Fair, and Poor. The correlations of these tests with scholarship are almost as high as that found for the Army Examination Alpha in the case of high school students²² and compare favorably with results of other studies among which may be named one made by Proctor.²³

²² Psychological Examining in United States Army, *Memoirs of the National Academy of Sciences*, Vol. XV. 1921.

²³ Proctor, W. M., *Psychological tests and the Probable School Success of High School Pupils*, J. Ed. Research, Vol. 1, pp. 258-270, Apr. 1920.

TABLE XI

Showing Correlation with Scholarship. (71 Cases, Grade 1 high to 3 high)

	r	P.E.
Non-Verbal 2	.413	.065
Haggerty Delta 1	.419	.064
Pressy Primer Scale	.322	.071
Haggerty Sigma 1	.722	.039

It was worth while also to find the closeness of relation between Non-Verbal 2 and the three other examinations used in this study. The coefficients in Table XI demonstrate that these tests do to a large degree measure the same abilities as Non-Verbal 2, this being especially true in the case of Delta 1. The closeness of the relation with the reading test Sigma 1 is also very interesting.

TABLE XII

Showing Correlation Between Primary Examinations. (71 Cases, Grades 1 high to 3 high)

	r	P.E.
Non-Verbal 2 and Delta 1	.726	.037
Non-Verbal 2 and Sigma 1	.695	.040
Non-Verbal 2 and Pressy	.616	.048

TENTATIVE STANDARDS.

The norms given below are derived from scores obtained in St. Paul, Minnesota. At present the examination is being given to pupils in cities and towns throughout several states in the Middle West, principally in Minnesota. When this testing is completed, representative norms will have been established for typical schools of the Middle West.

TABLE XIII

Giving Grade Medians, and Age Medians²⁴

	1B	1A	2B	2A	3B	3A	7	8	9
Oral Directions	4	7	8	9	10	11	7	8	9
Designs	5	6	7	9	9	11	8	9	10
Substitution (Div. by 3)	3	7	9	10	11	16	7	10	11
Pictorial Comp.	6	8	9	9	10	9	9	9	10
Pictorial Ident.	2	5	8	9	10	10	6	9	10
Total Score	19	34	42	47	50	55	36	45	50
No. of Pupils	139	181	112	154	117	76	266	263	118

²⁴ Age 7 includes all pupils who are between 6 years 6 months and 7 years 5 months. Ages 8 and 9 were figured by the same method.

NEWS AND COMMENT.

In this section the editors will attempt to present in brief form any news concerning important movements in the field of Applied Psychology which are taking place in the world; also reviews of the results of significant investigations in the field not published in the Journal. Brief and terse reports from contributing editors and others bearing on these points will therefore be gladly received and published so far as their merit and space will permit. We are particularly desirous of keeping our readers informed in regard to the more important developments in this field.

NEW BUREAU OF PERSONNEL ADMINISTRATION

To study and formulate the best methods of selection, transfer and promotion of public service employees is the primary function of the newly-organized Bureau of Personnel Administration at Washington, D. C. Professor L. L. Thurstone, head of the Department of Education and Psychology at Carnegie Institute of Technology, Pittsburgh, has been appointed Director of Research in charge of the Bureau, effective January 1st, 1923.

The Bureau is to be affiliated with the Institute of Government Research, and has been privately endowed subsequent to appeals for its establishment from the United States Civil Service Commission. An advisory board of five members of the national civil service commissions will supervise the operation of the bureau.

Mr. Thurstone and his associates will immediately take up the study of the selection, transfer and promotion on a basis of merit of all public service employees: classification of the positions in public service; methods of rating efficiency; reorganization of training courses for federal employes; recommendation of postal clerk and carrier tests; organization and procedure of district offices of the Civil Service Commission; and various other personnel problems outlined by the advisory board. Its field of work covers investigations of state and municipal employee problems in addition to those of the federal service.

SECOND INTERNATIONAL CONFERENCE OF PSYCHO-
TECHNICS¹ APPLIED TO VOCATIONAL GUIDANCE
AND TO SCIENTIFIC MANAGEMENT.

BY HARRY D. KITSON
Indiana University

European psychologists have made two correlations that American psychologists have not yet sufficiently established: (1) that between their science and scientific management; (2) that between the psychological problems of scientific management and those of vocational guidance. European psychologists see these relations clearly, as is evidenced by the fact that they have for three years held annual international conferences for the application of psychology to these two fields.

The first of these conferences was held at Geneva in 1920, the second at Barcelona in 1921, and the third at Milan, in 1922. The complete report of the second is now in the hands of the writer, and in fulfillment of a promise made in a previous number of this Journal, a resume of the proceedings is herewith presented.²

The conference was organized in three sections corresponding to the three main topics of discussion:

Section I. Vocational Selection and Study of Work.

- A. Vocational selection: General methods of research in psychotechnics applied to vocational selection. Inventory of the vocations already studied by psycho-technicians, examination of their results with a view to their application in vocational guidance. Program of international investigation in the determination of aptitudes required by certain vocations. Agreement upon a method for the investigation of the most important function in common occupations: motor aptitudes, attention, memory, etc., to the end of obtaining comparable results. Some special psychotechnical facts: Monotony, visual defects, tapping test.
- B. Researches on the adaptation of movements to vocational ends.

¹Psycho-technics, it will be remembered, is the term employed by Muensterberg (*Psychology and Industrial Efficiency*, Boston. Houghton Mifflin Co., 1913, p. 17) to denote the application of experimental psychology in economic fields. It is widely used in Europe.

²The papers are published as delivered in the original languages, by *L'Institut D'Orientacio Professional*, Barcelona, Sant Honorat, 5.

C. Study of vocational fatigue.

Methods for the discovery of objective signs of physical and mental fatigue.

Fatigue and Wages.

Fatigueability and Vocational Usefulness.

Section II. Vocational Guidance in General.

The scientific nature and the essential functions of Vocational Guidance.

What information should be called for on the medical, pedagogical and economic record blanks of the child?

How should these facts be secured; how used?

The problem of moral aptitudes.

Section III. Psychotechnics applied to vocational guidance.

What tests should be employed for the psychophysiological examination of those being guided?

What scientific facts are available for a classification of jobs according to the aptitudes which they demand of workers?

How can one evaluate the results of the mental examination?

It is manifestly impossible to present here a complete report of the discussions, but an idea of their nature may be secured from a list of the forty-six papers read. A star preceding a topic indicates that the delegate was present and read his paper in person.

*The Service of the School in the Determination of Aptitudes, by Otto Lipmann, director of Institute of Applied Psychology, Berlin.

*Contribution of the School in the Investigation of Aptitudes, How to Obtain Them and How to Utilize Them, by J. Ruiz Castella, director of Institute of Vocational Guidance, Barcelona.

*Some Results of Vocational Psychotechnic, by Walter Moede, professor at the Technische Hochschule, Charlottenburg, Berlin.

*The Principal Problems of Vocational Guidance, by Claparede, University of Geneva.

*Motor Tests in Vocational Guidance, by A. G. Christiaens, Director of the Inter-Communal Bureau of Vocational Guidance, Brussels.

*Tests of Aptitude in Drawing and their Application in Vocational Guidance, by O. Decroly, professor of Child Psychology in the Free University of Brussels.

*Psychological Automatism and Vocational Guidance, by Geo. Dweishauvers, Professor of Normal Studies and Director of Laboratory of Experimental Psychology, Barcelona.

The Role of Rhythm in Vocational Work, by M. J. Waldsburger, collaborator in the Laboratory of Experimental Psychology, Ecole Pratique des Hautes Etudes, Paris.

*The Application of Physiology to the Arrangement of Work, by Marcel Frois, Professor of Hygiene, Institut Lannelongue d'Hygiene Sociale, Paris.

The Economy of Human Effort in Industry, by Eric Farmer, Investigator to the Industrial Fatigue Research Board, London.

*Vocational Guidance for Industrial Occupations, by Geo. H. Miles, D. Sc., Secretary of the National Institute of Industrial Psychology, London.

The Investigation of the Motor Function, by Raymond Buysse, Director of Primary Education, Saint Gilles, Belgium.

Two Communications from the International Bureau of Labor: 1. The Scientific Selection of Personnel in the United States of America. 2. The International Bureau of Labor and Vocational Guidance.

Scientific Management and Mental Hygiene, by Genil-Perrin, head of the clinic of mental disease in the School of Medicine of Paris, General Secretary of the Mental Hygiene Association of France.

A Few Suggestions, by A. Imbert, Honorary Professor in the School of Medi-

- cine, Montpellier; Professor of the Physiology of Work in the School of Medicine Marseilles.
- *The School, the Laboratory, the Factory and the Development of Manual Skill, by J. M. Laby, of the Laboratory of Experimental Psychology of the Sorbonne, and of the Association of Mental Hygiene, Paris.
Synopsis of Investigation on Vocational Ability in Engineering Trades, by Max Tagg, Department of Engineering, Acton & Chiswick Polytechnic, Bedford Park, London.
 - *Some Socio-Economic Aspects of Scientific Management, by J. M. Tallada, Engineer, member of the Advisory Council of the Institute of Vocational Guidance, Barcelona.
 - *The Medico-Anthropometric Record Blank, by L. Trias de Bes, head of the Medico-Anthropometric Laboratory of the Vocational Guidance Institute, Barcelona.
 - *Uses of Medical Measurements in the Vocational Guidance Institute, by J. M. Lucena, Assistant in the Medico-Anthropometric laboratory of the Vocational Guidance Institute, Barcelona.
 - *Immediate Organization of the Vocational Guidance of Youth, by E. Gauthier, head of the Office Regional de la Main d'Oeuvre du Ministère de Travail de France, Paris.
How to Make Vocational Guidance Scientific, by Richard Liebenberg, Employment Director of the City of Berlin.
 - *Vocational Guidance for Deaf Mutes, by P. Barnils, Director of the School for Deaf and Mutes, Vilajoana (Spain.)
Psychotechnic in Italy, by G. C. Ferrari, Bologna.
Establishment of Vocational Guidance Offices by F. Mauvezin, Director, Chambre de Metiers of the Gironde and of the South-West Boireaux.
 - The Berlin "Classes for Gifted Children" After Four Years Trial, by Kurt Piorkowski, Director of the Orga-Institut, Berlin.
 - *Vocational Guidance as Practiced at the Institut, Emil Metz, by Aloyse Robert, of the Institut Emil Metz, Dommeldange (Luxemburg).
 - *Essentials of Vocational Guidance, by J. Ruiz-Castella, Director of the Institute of Vocational Guidance, Barcelona.
 - *Scientific Relations Between the Medical Inspection in Schools and the Institute of Vocational Guidance, by Manuel Salvat Espasa, Medical Inspector of Schools, Barcelona.
Necessity of a Practical Method of Vocational Guidance, by R. Thery, head of vocational guidance office at Nantes.
 - *The Participation of the Placement Office of the Department of the Seine in the Work of Vocational Guidance, by A. Touzaa, Bureau of Labor Statistics of the Seine, Paris.
 - *A Test of "Sympathetic Intuition" for the Examination of Moral Aptitudes, by Franziska Baumgarten, Berlin.
 - *Cario-Vascular Accompaniments of Mental Work, by E. Mira, head of the Psycho-Technic Laboratory of the Vocational Guidance Institute, Barcelona.
 - *The Psycho-Physiological Examination in the Interest of Education and Vocational Guidance, by N. Braunhausen, Professor of Experimental Psychology, Lycee of Luxembourg.
 - *The Psycho-Galvanic Phenomenon in Psycho-Technics, by J. M. Bellido, Assistant Director of the Institute of Physiology, Barcelona.
General Ability and the Objective Psychic Function, by Julius de Boer, of Nimegue (Holland).
 - English Contributions to the Study of Mental Tests, by Cyril Burt, Psychologist to the London County Council.
 - *The Measurement of the Intelligence of School Children in Barcelona, by Llorens Cabos, Principal of the National Schools of Barcelona.
The Intelligence Aspect of Arithmetic, by D. J. Collar, London.
The Place of the Psychologist in Industry, by Frank B. and Lillian M. Gilbreth, Montclair, New Jersey, U. S. A.
 - *Psychotechnical Study of Watchmaking, by Hugo Heinis, Laboratory of Psychotechnics of the University of Geneva.
The Determination of Vocational Aptitude, by A. Imbert, Professor of the Physiology of Work at the School of Medicine, Marseilles.
 - *The Binet Tests and the American Revision, by J. M. Lahy, and A. Fessard, Paris.
 - *Education and Adaptability, by J. M. Lucena, Assistant in the Psychometric Laboratory of the Vocational Guidance Institute, Barcelona.
 - *Tests for Bridgebuilders, by Aloyse Robert, of the Institut Emil Metz, Dommeldange, (Luxembourg.)
 - *The Relation Between Concentrated Attention and the Extent of the Visual Field, by C. Soler Dopff, Assistant at the Psychometric laboratory, Vocational Guidance Institute, Barcelona.

*Psychotechnical Methods employed in the Laboratory of the City of Amsterdam, by G. van Wayenburg, Privat-Docent of Psychology and Paidology, University of Amsterdam.

Questions that were treated in extended discussion are: Shall the schools assume chief responsibility for vocational guidance?

At what age shall formal guidance begin? Shall the guidance be continuous, as proposed by some in the United States, or shall it be given once for all? How many times shall tests be given for the determination of aptitude? After warm discussion of these questions in sectional and general conference, a series of conclusions and recommendations were adopted as representing the view of the majority of the delegates:

I. Regarding the practice of vocational guidance:

1. Vocational guidance in its scientific aspect is an economic and a social problem. In consequence, it cannot be narrowed down merely to a function of the schools; it must be attached to bureaus of placement and institutions insuring against unemployment.

2. A single examination of a child can give only approximate indications in a certain number of cases, and for a certain number of aptitudes; it would be desirable to extend this examination over a period of time before the child leaves school, and while he is in the first stages of apprenticeship. The special conditions of each bureau of vocational guidance should determine the extent of these examinations.

3. To the end of perfecting methods, the results which have already been obtained by persons who have been placed in industry should be studied scientifically, from the following points of view: technical progress, interest in work, fatigueability, and wages.

4. The medical record should be standardized and should contain two parts:

- a. One page containing observations of an anthropometric and physiopathological sort. This must be regarded as confidential.
- b. A record to be filed with the bureau of vocational guidance showing individual data which will permit the investigation of a group of occupations.

II. Regarding Methods of Increasing Industrial Efficiency:

1. Public agencies should favor physiological and psychological researches looking toward the establishment of scientific management and vocational guidance upon a scientific basis.

2. The conference recommends that governments subsidize laboratories working upon the standardization of tests of vocational aptitude, and that the results of the investigation in the laboratories of the different countries be combined.

3. In the light of numerous sociological and physiological data it is recommended that eight hours of effective work be considered as a normal working day in the majority of industrial occupations.

4. The conference recommends that among the agencies of transportation (railroads, street cars, automobiles, steam ships) workers shall not be hired except upon psychological examination for vocational aptitude. Psychiatric control should also be exercised over such workers.

5. It is desirable to define more accurately such terms as: practical intelligence, mechanical aptitude, higher, medium and lower occupations; acquired, formal and material aptitudes.

One of the significant features on the program was the voice of the International Bureau of Labor, which appeared twice, once with an assurance of its moral support; again, with a statement regarding the administration of employment management in American industrial establishments. The Bureau, established in connection with the League of Nations, and with headquarters at Geneva, is deeply interested in vocational guidance. Its pronouncement at this conference was so significant that it is herewith translated in full:

"The International Bureau of Labor and Vocational Guidance."

"Vocational guidance is a problem that interests directly the activity of the International Bureau of Labor because it constitutes one of the primary conditions for the realization of the program of reforms set forth in the preamble of Part XIII of the treaty of Versailles. It may be considered as the basis of practically all the measures prescribed in the treaty of peace for the amelioration of the conditions of work and for the realization of social justice, which must form the bases of universal peace.

Specifically, it plays a capital role in 'the organization of technical and vocational instruction,' because it has for its purpose the selection of the well-endowed. It is the condition of 'rational recruitment of labor,' because it seeks to determine exactly the aptitudes of each one. It furnishes the worker with a weapon in his 'battle against unemployment,' because it increases his technical value. It aids in assuring him a 'living wage' because it raises his earning capacity to a maximum. It constitutes a 'protection to workers against general or occupational diseases and against industrial accidents, because it tends to give each one a job that accords with his physical capacities. It is a measure of 'protection for children and adolescents', because on the one hand it has for its purpose to guide them at the threshold of life; and on the other, to determine 'the conditions necessary for the extension of their education and for the development of their physique,' it has an important bearing upon 'the protection of women,' because it takes account of the physical and moral factors involved in women's work. It plays a part in that which concerns the disabled veterans, because it is at the basis of vocational re-education, the results of which will serve to determine the degree of in-

capacity of the *mutilé*. It can contribute to the 'defense of workers employed in foreign lands' because the phenomenon of emigration requires a scientific study of the ethnic elements whose influence from a vocational point of view is clearly important. Finally, having for its ideal the realization for each individual of the occupation which best suits his aptitudes and his tastes, and which procures for him the deepest and most lasting satisfactions, vocational guidance appears to be that condition of 'moral and intellectual well-being of wage-earners', recognized by the signatory powers of the Treaty of Peace as being 'of primary importance from an international point of view'. Accordingly, the analysis of capacities is the foundation of a scientific organization of work and of an amelioration of the conditions of work.

The International Bureau of Labor, whose functions comprize the centralization and distribution of facts concerning the international standards of working conditions, has, since its inception, devoted special attention to the question of vocational guidance. It has already organized services of emigration, of unemployment, and of industrial hygiene, in all of which it has recognized the important part played by vocational guidance. Similarly, in its efforts on behalf of those disabled in the war and in industry, it has been obliged to consider vocational guidance, notably in an article devoted to the vocational re-education of veterans in Great Britain. If the International Bureau of Labor has not been able heretofore to realize the establishment of a technical service devoted exclusively to the study of questions relative to education and apprenticeship, it has nevertheless charged one of its collaborators with the duty of centralizing all facts relating to these questions, of following their developments, and of maintaining contact with the measures taken or the experiments tried in the different countries.

The information service of the bureau takes great pains to collate from different journals and periodicals all facts bearing upon this question. A special category is devoted to technical education in the 'Facts of the Day' published bi-weekly. In addition, the Bureau proposes to publish in the 'International Labor Review' special articles upon vocational guidance written by specialists who have promised their cooperation.

Animated by the desire to follow as closely as possible the development of vocational guidance and cognate problems, the Bureau is constantly enlarging its library so as to include the most important works upon the question. It announces regularly in its 'bibliographical notes' recent publications on vocational education. In addition, it publishes in the International Labor Review bibliographical notices of publications pertaining to the general subject.

Finally, it has already entered into relation with such organizations as The Taylor Society of the United States the Industrial Fatigue Board of Great Britain, and the *Cours d'Orientation Professionnelle de France*, as well as with individuals who are devoting their attention to the problems of vocational guidance. The department of public instruction of the canton of Geneva, in particular, has willingly placed at its disposal results of valuable experiments carried on in cantons of Switzerland.

The International Bureau of Labor followed with the greatest interest the deliberations of the first international conference of psycho-technics applied to vocational guidance, which was held at Geneva in 1920, under the presidency of Professor Claparède; and M. Albert Thomas, in greeting the delegates who came to visit the International Bureau of Labor, assured them warmly that he considered their contributions most important for the realization of the tasks which the bureau has undertaken.

Desirous of neglecting no aspect of the task before it, and convinced of the immense importance that vocational guidance will assume in the scientific organization of work, the International Bureau of Labor will be particularly happy to follow the deliberations and to heed the counsels of the Second International Conference of Psycho-technics Applied to Vocational Guidance and Scientific Management."

The rule that delegate representation in the International Bureau of Labor is contingent upon membership in the League of Nations, does not at all bar American psychologists who are applying their science in the field of labor problems, from associating themselves with the scientific movements sponsored by the Bureau. They are cordially urged to read the publication of the Bureau—the International Labor Review—and to send reports of their investigations to the office of the psycho-technical representative of the Bureau—the Institute J. J. Rosseau, Geneva.

Space forbids detailed treatment of the other important events of the conference. The evidence here presented is sufficient, however, to show that European psychologists are seeing clearly the close relations that exist between psychology and scientific management and between the latter and vocational guidance, and that they are seriously attacking the salient problems in both fields.

PROGRAM AND WORK OF THE AMERICAN NATIONAL VOCATIONAL GUIDANCE ASSOCIATION.

The National Vocational Guidance Association of the United States of America met in annual convention at Detroit, Michigan, U. S. A., December 1 and 2. This organization has a special Section on Vocational Placement, on Vocational Guidance in the High School. A General and a Psychological Section. At the Psychological Section papers were presented on the following subjects:

1. Relations of Mental Levels to Industrial Life, by Helen T. Woolley, Assistant Director of the Merrill-Palmer School, Detroit.
2. An Investigation of the Vocational Changes of a Thousand Eminent Men, by Harry D. Kitson, Professor of Psychology, Indiana University.
3. The Use of Mental Tests in Educational and Vocational Guidance in fourteen Cities, by Mary Hays, Childrens Bureau, Washington, D. C.
4. Principles in Technique of Vocational Guidance found by the Institute of Educational Research, by Herbert A. Toops, Columbia University.
5. An Experiment in Educational Guidance, by Ruth Clark, Vocational Service for Juniors, New York City.

The Association voted to enter into collaboration with the Vocation Guidance Bureau of Harvard University in the publication of the official Bulletin of the Association. The officers of the Association for the coming year are Harry D. Kitson, Indiana University, President; John Brewer, Harvard University, secretary, and Arthur F. Payne, Harvard University, Treasurer.

MENTAL GROWTH CURVE OF NORMAL AND SUPERIOR CHILDREN STUDIED BY MEANS OF CONSECUTIVE INTELLIGENCE EXAMINATIONS.

This study, made by B. J. Baldwin and Lorle I. Stecher, and published in the University of Iowa Studies in Child Welfare, as Vol. II, No. 1, January, 1922. Pp. 1-61, is the first serious effort to determine the nature of the mental growth curve of superior children as compared with that of average children by Binet tests repeated on the same children through a number of years. 143 children figure in the study, and 485 separate examinations were made, by four different examiners. The I. Q.'s ranged roughly from 90 to 160. The initial examinations were made at the Iowa Child Welfare Research Station in 1917. The time intervals between successive examinations varied from a few months to over two years, and the number of examinations per child varied from two to five. Terman's revision of the tests were used throughout.

The results are presented under three main headings: (1) the mental growth curve; (2) the possibility of prediction of mental growth; (3) the relation between physical and mental growth. As the examinations did not occur on or near the birthdays of the children, the mental ages for the exact chronological ages of 5, 6, 7, etc., were computed on the basis of the rate of mental progress between successive examinations. To get the mental growth curve these resultant mental ages were averaged for each chronological age from 5 to 14. The results are given separately throughout for boys and girls, and for the two grades of intelligence as given by the I. Q., termed "average" and superior." I. Q.'s for the average ranged from 90 to 110, and for the superior 111 and

over. The main conclusions reached from this first part of the study are (1) that the mental growth curves of average and superior children diverge, the difference in mental age between the two classes increasing with age; (2) that there is a significant change in the trend of the curve on approach to adolescence, and that an adolescent superiority of girls is also revealed; (3) that the general prepubertal increase appears earlier in the superior children; (4) that the I. Q. curves are approximately horizontal, and the I. Q. is constant within limitations; (5) that the increase in the mean I. Q. with an increase in age that was found is probably an effect of a greater habituation or practice in passing the tests.

Considerable variation in successive I. Q.'s of the same child were found frequently, leading the authors to give the caution that "extreme care should be exercised about making any dogmatic statements in regard to what a child's future status will be." The deviation of each child's I. Q. on each particular examination from his average I. Q. for all his examinations was computed. The mean I. Q. deviations from these averages were $4.48 \pm .53$ for superior boys, $6.92 \pm .78$ for superior girls, $2.65 \pm .23$ for average boys, and $3.80 \pm .29$ for average girls. The frequency of positive changes in the I. Q., that is, increases, is greater than the negative changes. This preponderance of positive changes is more marked for superior than for average children. Pearson co-efficients of correlation between the I. Q.'s of successive examinations range from $+.69 \pm .05$, in the case of the first and third examinations, to $+.93 \pm .02$, in the case of the third and fourth examinations. There is no tendency for the correlation to decrease with increase in the interval between examinations. The more important conclusions summarized for this second section of the results are as follows: (1) There is a tendency for girls to vary more than boys, for superior to vary more than average children, and for older to vary more than younger. (2) between the first and second examinations most cases show an I. Q. change of less than five points; between the first and later examinations more cases show larger changes, with the increases in I. Q. becoming more numerous. (3) The greater tendency for superior children to show increases in I. Q. on later examinations, as compared with average children, is probably due to the former profiting more readily by practice. (4) "No final determination can be made of the effect of chronological age or of the interval between examinations on the change in I. Q.'s."

Height and weight measurements were made, also X-ray photos of the carpal bones. The mean mental age of the physiologically accelerated children was found uniformly higher than that of the

physiologically retarded. Correlations found were as follows: Height and mental age, boys $+.84 \pm .02$, for girls $+.89 \pm .02$. Weight and mental age, boys, $+.86 \pm .02$, girls, $+.77 \pm .04$. Mental age and X-ray result (exposed area of carpal bones of right wrist), boys, $+.87 \pm .02$, girls, $+.87 \pm .02$.

In their general conclusion the authors note that "the use of the Stanford Revision of the Binet scale even for the relatively short period of four years shows the unsuitability of this scale in its present form as a means for measuring mental growth. The limited number of alternative tests results in a certain practice effect on repeated examinations. Another defect of the present system of tests is in the lack of a sufficient number of tests at the higher levels to measure the mental growth that apparently goes on in a bright young child even after the exhaustion of the 16 or 18 year old tests."

Their conclusion that the results do not decide the question of a general tendency of the I. Q. of superior children to change with increasing age, because of the inadequacy of the tests is disappointing. The reviewer believes that this conclusion is, in part, the result of an error in the method of treating the data, and, in part, to their failure to consider other interpretations that explain more of the facts as we now have them. They are undoubtedly correct in attributing some of the gain in I. Q. with increasing age to practice effect. They are dealing largely with the upper levels of the scale where the tests are most affected by this fault. It is probable also that a small University town, with the children figuring in the study coming from the University school, and their parents undoubtedly more informed about the tests than would have been true of any other group, resulted in an unusual amount of coaching in the tests. But that this can explain all the increase in I. Q. that the results really show is highly improbable. Since they make it the sole explanation of one of the most important parts of the study, it seems that they might have made a greater effort to treat the results in a way that would support this explanation. But little attempt is made to do this. Their method of treating the results does not show the amount of increase in I. Q. with increasing age, but assumes a constancy of the I. Q. in the first place. The children are divided at each age into two classes, those below an I. Q. of 111 the average, and those above an I. Q. of 110, the superior. For each class the mean mental age, with I. Q. derived from this, was then computed at each age. If the I. Q. did increase with age, then obviously many of the older children classed as superior would have classed as average when younger, and many of the younger children now classed as average

would have been classed as superior later. This results in a raising of the mean I. Q. for the younger children and a lowering of the mean I. Q. for the older children. The authors give the raw data, including age, M. A. and I. Q. of each child at each examination. By summing up all the age intervals between examinations and also the I. Q. changes, irrespective of age, the reviewer finds the following general tendency of the I. Q. to increase with age. Average yearly gain for superior boys, 3.89 points, for superior girls, 3.92 points, for average boys and girls together .45 points. Of those classes as average, 37 had I. Q.'s above 100, and 15 had I. Q.'s below 100. The class in which a child belongs, in each case, is determined by his average I. Q. for the two to five examinations given him. This gives an entirely different aspect to the results on this point. The author's analysis shows a slight increase; this reveals a very marked increase.

In my results on the re-examination of subnormal children, which showed a decrease in I. Q. with age, I attributed this fact to a decreasing rate with age of mental development of normal children, when this development is measured in terms of absolute, real units of development, and I anticipated that the I. Q. of superior children would for the same reason increase with age and explained how this would come about. (See *Journal Applied Psychology*, September, 1921). The present Iowa results on this point seem to me chiefly an empirical proof of this deduction.

Director State Research Bureau
St. Paul, Minn.

DR. F. KUHLMANN.

BOOK REVIEWS.

ALBERT S. OSBORN. *The Problem of Proof*. Matthew Bender Bender & Company, New York, 1922, pp. 526, Introduction by Professor J. H. Wigmore.

Osborn—whose *questioned documents* is the classic volume on the handling of disputed writing with the aid of modern instruments and in the light of scientific analysis—in his new book *The Problem of Proof*, considers the manner in which facts should be collected in preparation for a trial in which identification of handwriting or typewriting is concerned and how the evidence obtained should be presented. Trial lawyers, handwriting experts, and judges should read the book with care. Much practical and very valuable advice is given as to how to obtain writing standards with which to compare disputed material, how writing exhibits should be handled and filed and how shown in court. The attitude that the expert should assume while on the witness stand is discussed at length. A most interesting series of outlines to be used in the cross examination of incompetent and corrupt witnesses is given. And there are some candid comments on the designing and lighting of court rooms and hold-overs in court procedure that should be pondered by all those who occupy positions of authority in the legal realm.

All students concerned with problems of handwriting should know Osborn's work. For the general psychologist, in the particular book under consideration, the chapter on "Form Blindness or Defects of Sight in Relation to the Problem of Proof" is of most interest. It is a revised statement of an article published some years since by Osborn and cites experiments to show that something called form-blindness operates to a high degree in many observers and makes it very difficult for them to pass accurate judgements on writing or to identify hands except when greatly enlarged.

The psychologist should also profit by the discussion of handwriting recognition and identification by the untrained man and the analysis of what "knowing a hand" really means. There are problems raised that should be carried into the laboratory. The clear-cut description of the characteristics of natural writing and of the inherent evidences of genuineness or the reverse are important. The practical psychologist will read with interest the chapter on persuasion and the student of applied psychology find advice as to the course to follow in analysis of anonymous letters and in attempts to identify the writers of them.

University of Wyoming

JUNE E. DOWNEY

GUY THOMAS BUSWELL, *Fundamental Reading Habits; A Study of Their Development; Supplementary Educational Monograph No. 21*, Chicago, Department of Education, The University of Chicago, 1922, pp. 14+150. \$1.50.

The purpose of this study is to detect the fundamental reading habits of children and to compare them with the mature reading habits of adults. The assumption of the author is that by knowing what the child's early reading habits are to begin with, and what the mature habits of good readers are as an end product, we shall be better able to direct the child toward the desired goal of efficient reading. By knowing the route and the individual difficulties to be met the teacher is enabled to adapt the plan to the need.

Since all satisfactory methods of studying reading habits must rest on either a direct or an indirect measurement of some aspect of the reading process, the author has selected photographic reproductions of eye movements as the most objective in character, and the least subjective in judgment. The apparatus used was an improved type of that used by Dr. C. T. Gray in his studies in 1917. Dictaphone records were taken alongside the photographic reproductions of eye movements and thus a most thorough comparison of the two sets of data is made possible.

The study analyzes the reading of 187 different subjects ranging from children in grade I-B of the elementary school to adult college students. Frequent tables and charts show the relative stages of development for many of the characteristic reading habits. Probably the most striking feature of the data shown is the clearly demonstrated importance of the first four years of reading upon the development of mature reading habits. Another marked change in reading habits occurs in the early years in High School. The growth curves are very much alike for both oral and silent reading, though silent reading develops maturity in reading habits slightly faster than oral reading.

Dr. Buswell devotes about one third of his monograph to a detailed analysis of first grade reading. The children studied came from two different groups each of which had been taught by a radically different method. The analysis brings out the fundamental weaknesses of each method. The data is presented in exceptionally complete form and includes charts showing the number and extent of eye movements, the time of each fixation, the dictaphone reproductions of the reading, and the rhythmic expression of the subject while reading. This part of the monograph should prove especially valuable to teachers in that it will enable them to make a better analysis of the fundamentals of the methods they are employing, and should assist them greatly in placing a truer evaluation upon their own work. The latter third of the book is concerned with individual variations in reading habits and with the results of specific training. Characteristic cases from the 2nd, 3rd, 4th and 5th grades, and from High School and College are described and analyzed.

It is not assumed that a study of eye movements alone will make possible a complete diagnosis of all reading difficulties, but the author insists that such study must be a part of every complete study of reading defects. It reveals many specific difficulties which demand specific treatment. Dr. Buswell is not in agreement with those who insist that no artificial devices shall be allowed a place

with remedial teaching because they may lack social values, or permanent purpose outside the schoolroom. He urges the use of any kind of artificial device which by its temporary use will remedy the child's reading defect or bring him back to normal more rapidly than by some less direct method. For training the child in correct habits of eye movement an exposure apparatus to be used in conjunction with the stereopticon is described. How successful this device may be is neither demonstrated nor forecasted; but, being neither too intricate in structure nor over expensive in manufacture, it should find a place in any laboratory devoted to experimental work in reading.

Ohio University

C. H. Growdon.

WILLIAM SCOTT GRAY, *Remedial Cases in Reading: Their Diagnosis and Treatment*—University of Chicago, 1922, Pp.vii-208. \$1.75.

The content of this monograph is well described by its title. It promises to be a very suggestive and helpful volume for superintendents, principals, supervisors, and teachers—whether regular or special. In these days of frequent measurements of all kinds there is sometimes a tendency to make hasty diagnosis. This monograph shows the need for very careful and definite study of each case before final diagnosis is made; it gives detailed information as to existing defects in each case, and describes the procedure which was followed in the attempt to remedy each defect.

The author holds that diagnosis of one's reading defects is insufficient evidence upon which to base treatment. The causes must be determined. Among the more common fundamental causes of defective reading are listed: inferior intelligence, congenital word blindness, poor auditory memory, defective vision, narrow span of recognition, ineffective eye movements, inadequate training in phonetics, inadequate attention to content, inadequate speaking and meaning vocabularies, speech defects, lack of interest, guessing, and timidity.

Five chapters of the book (150 pages) are devoted to a detailed report of 26 cases which came under the observation of the author or his immediate assistants. These cases are grouped into five typical classes as follows: (a) pupils who had made little or no progress in reading, (b) pupils who had encountered serious difficulties in interpretation, (c) pupils who encountered difficulties primarily in the mechanics of reading, (d) pupils who encountered difficulties in the rate of silent reading, and (e) pupils who had made progress in reading but were weak in practically all phases of reading. The latter chapter of the work deals with the organization of diagnostic work in the city of Toledo, Ohio.

The book neither attempts to cover all types of reading defects nor does it dogmatize on the remedial treatments to be applied. It is merely a plain statement of a considerable number of typical cases with a description of the remedial measures attempted, and with an attempt to evaluate the efficiency of the methods used. As a source of helpful suggestions for those who must handle reading defects it is worthy of a careful reading. We believe that the stimulus it will give toward the organization of groups interested in the study and treatment of reading defects, and the encouragement it gives to teachers who daily face reading difficulties make its appearance especially timely.

Ohio University

C. H. Growdon.

JOHN DEWEY, *Human Nature and Conduct*. New York, HENRY HOLT AND COMPANY, 1922.—pp. 336.

This late book from the facile pen of Professor Dewey gives to his many readers in his usual fresh and lucid style a constructive treatment of the primary facts of social psychology. It is the result of the re-writing and expansion of three lectures delivered by its author in Leland Stanford Junior University in the spring of 1918. While not purporting to be a treatment of social psychology, it nevertheless makes habit the central fact for the explanation of all social phenomena and seriously maintains that an understanding of habit is the key to social psychology.

The work is conveniently arranged in four parts as follows: First, The Place of Habit in Conduct; second, The Place of Impulse in Conduct; Third, The Place of Intelligence in Conduct; fourth Conclusion. Its fundamental premise is that all nature, including man and society, is continuous. Human conduct is explained as the result of the interaction between the instinctive elements in man's nature and the factors of the natural and social environment. Differing from those writers on social psychology and the social sciences who make some subjective factor, instinct or impulse, or even intelligence, as the central point of reference, and, differing from those who stress exclusively the forces of an environment as determinants of conduct, Professor Dewey takes a middle course and holds that there are forces *in* man as well as *without* him. "Mind," he holds, "can be understood in the concrete only as a system of beliefs, desires and purposes which are formed in the interaction of biological aptitudes with a social environment." Yet throughout the author contends that the primary facts of social psychology center about habit or custom, and by so doing he gives to the subjective factors a position of relative frailty as compared with the exterior forces of the environment. "Impulses, although first in time, are never primary in fact; they are secondary and dependent," simply because they are aroused and given meaning, moulded and developed by a relatively fixed medium. Impulses, however, are important as "organs or reorganization and readjustment" when habits fail to meet the demands of life; they are pivots "for giving new direction to old habits and changing their quality." The instincts of man being plastic and malleable, as witnessed by the wide diversity of social institutions and moral codes built upon them, the wise student of social reform will "need to find out just how different customs shape the desires, beliefs, purposes of those who are affected by them." This is the crucial problem of social psychology.

.....
Professor Dewey has here employed his broad knowledge and ripened judgment to a critical examination of the foundation facts of all the social sciences. Though not everyone will be able to agree with his findings, certainly no student of social psychology and social philosophy concerned to get at the facts can fail to profit from a careful reading of this original and penetrating analysis of human behavior.

Ohio University

Walter S. Gamertsfelder

- Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio, or William F. Book, Indiana University, Bloomington, Indiana.
- BARKER, ELSA. *Fielding Sargent*. E. P. Dutton & Company, 681 Fifth Avenue, New York. 319 pp.
- BONNER, H. R. *Statistics of City School Systems*. Department of the Interior, Bulletin No. 17, 1922. Washington, D. C. 173 pp.
- COX, FLORENCE C. *Major Projects in Elementary Schools*. Department of the Interior, Bulletin No. 36, 1921. Washington, D. C. 43 pp.
- DELGADO, HONORIO F. (With Introduction by William A. White.) *Algunos Aspectos de la Psicología Del Niño*. Casa Editora Imprenta la Opinion Nacional, Lima, Peru. 82 pp.
- DUNLAP, KNIGHT. *Elements of Scientific Psychology*. C. V. Mosby Company, St. Louis, Mo. 362 pp.
- EGBERT, JAMES E. *University Summer Schools*. Department of the Interior, Bulletin No. 31, 1922. Washington, D. C. 14 pp.
- GATES, ARTHUR I. *The Psychology of Reading and Spelling*. Teachers' College Columbia University, New York City. 108 pp.
- GEBHART, JOHN C. *Malnutrition and School Feeding*. Department of the Interior, Bulletin No. 37, 1921. Washington, D. C. 39 pp.
- GREENE, ARTHUR M., JR. *Engineering Education After The War*. Department of the Interior, Bulletin No. 50, 1921. Washington, D. C. 27 pp.
- HOLLINGWORTH, L., *Judging Human Character...* D. Appleton & Company, Publishers, New York. 268 pp.
- JACKSON, HENRY E. *Robinson Crusoe—Social Engineer*. E. P. Dutton & Company, 681 Fifth Avenue, New York. 301 pp.
- KENNETH, JOHN H. *Osmics, The Science of Smell*. Oliver & Boyd, Edinburgh. 42 pp.
- MOORE, THOMAS VERNER. *Percy Bysshe Shelley*. Psychological Monographs. Psychological Review Company, Princeton, N. J. 62 pp.
- PIERON, HENRI *L'Annee Psychologique* (1920-1921) Librairie Felix Alcan, 108, Boulevard Saint-Germain, Paris. 608 pp.
- PRESCOTT, F. C. *The Poetic Mind*. The Macmillan Company, New York. 308 pp.
- Record of Current Educational Publications*. Bureau of Education, Department of the Interior, Bulletin No. 21, 1922. Washington, D. C. 42 pp.
- Segona Conferencia Internacional De Psicotecnica Aplicada a l'Orientacio Professional i a l'Organitzacio Cientifica del Treball*. Institut d'Orientacio Professional, Sant Honorat, 5, Barcelona. 419 pp.
- THOMPSON, J. DAVID. *Personnel Research Agencies*. Department of Labor. Bulletin No. 299, November, 1921. Washington, D. C. 207 pp.
- VALDIZAN, HERMILIO AND HONORIO F. DELGADO. *Revista de Psiquiatria y Disciplinas Conexas*. C. T. Southwell, Pando 765, Lima, Peru. 175 pp.
- WARREN, HOWARD C. *Elements of Human Psychology*. Houghton Mifflin Company, New York. 416 pp.
- ZOOK, GEORGE F. *The Residence of Students in Universities and Colleges*. Department of the Interior, Bulletin No. 18, 1922. Washington, D. C. 11 pp.

* Mention here does not preclude further comment.

INDEX OF SUBJECTS

(The subject headings of all contributions except Book Reviews and Notes are printed in SMALL CAPITALS.)

ADDING MACHINE, INTERCORRELATION OF TEST ON	172	FACTORS OF AESTHETIC JUDGMENT	34, 120	
ADVERTISING, DEVELOPMENT OF ART-FORMS IN MAGAZINE, 59; COLOR IN MAGAZINE, 64; NEGATIVE SUGGESTION IN,	66	FINDING EQUIVALENT SCORES IN TESTS OF INTELLIGENCE	29	
ADVERTISING APPEAL, THE TESTIMONIAL AS,	192	GENERAL INTELLIGENCE COMPARISON OF THREE TESTS OF	391	
AESTHETIC JUDGMENT, SOME FACTORS OF	34, 120	GROUP INTELLIGENCE EXAMINATION FOR PRIMARY CHILDREN	403	
ART-FORMS IN MAGAZINE ADVERTISING, DEVELOPMENT OF	59	FIRST GRADE, MASS MENTAL TEST FOR USE WITH KINDERGARTEN AND	89	
Character Analysis, Extent Used in Business World	302	HIGH SCHOOL COURSES AS MEASURED BY THE OTIS TEST	185	
CHILDREN, CLASSIFYING PRIMARY GRADE, BY MENTAL AGES, 276; COLOR PREFERENCES OF,	255	High School Seniors, Intelligence of	85	
CLASSIFYING PRIMARY GRADE CHILDREN BY MENTAL AGE, AN EXPERIMENT IN	276	HORMONES AND INTELLIGENCE	221	
COLOR PREFERENCES OF CHILDREN	255	Human Behavior (Paton)	215	
COMPARISON TESTS FOR INVESTIGATING THE IDEATIONAL CONTENT OF THE MORAL CONCEPTS	235	IDEATIONAL CONTENT OF THE MORAL CONCEPTS, COMPARISON TEST FOR INVESTIGATING THE	233	
CONCEPTS, COMPARISON TEST FOR INVESTIGATING THE IDEATIONAL CONTENT OF THE MORAL	235	Industry, Personnel Relations in (Simons)	307	
CONVERSION OF TEST SCORES INTO SERIES WHICH SHALL HAVE ANY ASSIGNED MEAN AND DEGREE OF DISPERSION	298	Intelligence, Morphological Aspect of (Naccarati), 303; of High School Seniors (Book)	95; QUOTIENT, FROM TWO VIEWPOINTS, 267; TESTS, FINDING EQUIVALENT SCORES IN, 29; HORMONES AND,	221
Developing Mental Power (Stratton)	216	INTERCORRELATIONS OF, TESTS ON ADDING MACHINE	172	
DISPERSION, CONVERSION OF TEST SCORES INTO SERIES WHICH WILL HAVE ANY ASSIGNED MEAN AND DEGREE OF	298	JUDGMENT, RELIABILITY OF, OF PERSONAL TRAITS, 161; SOME FACTORS OF AESTHETIC	34, 120	
EXPERIMENT IN CLASSIFYING PRIMARY GRADE CHILDREN BY MENTAL AGE	276	KINDERGARTEN AND FIRST GRADE, MASS MENTAL TEST FOR	89	
		MEASURES OF GENERAL INTELLIGENCE AS INDICES OF SUCCESS IN TRADE LEARNING	311	

INDEX OF SUBJECTS

Measuring the Results of Instruction in Colleges and Schools, An Efficient Method for (Book)	86	PRIMARY GRADE CHILDREN, CLASSIFIED BY MENTAL AGE	276
MENTAL AGE, CLASSIFYING PRIMARY GRADE CHILDREN BY	276	PSYCHOLOGY OF ADVERTISING, MINOR STUDIES IN, 59; of Selling Life Insurance, (Strong)	304
MENTAL FATIGUE OF MIXED AND FULL BLOOD INDIANS	331	Public Opinion (Lippman)	306
Mental Power, Developing	216	PUBLICITY, EFFECT OF LETTERS AND SYLLABLES IN	198
MENTAL SURVEYS, METHODS AND RESULTS OF	1	RELIABILITY OF JUDGMENT OF PERSONAL TRAITS	161
MENTAL TESTS AS AN AID IN THE ANALYSIS OF MENTAL CONSTITUTION	349	SELECTION OF A SUCCESSFUL SECRETARY	156
METHODS FOR STUDY OF VOCATIONAL INTERESTS, 243; AND RESULTS OF MENTAL SURVEYS	1	SOCIAL RELATIONS TEST	69
Mind and Work (Myers)	84	SPEECH AND WILL-TEMPERAMENT	291
Mind of the Buyer, the (Kitson)	215	STANDARDIZATION OF TAPPING TEST	113
MORAL CONCEPTS, COMPARISON TEST FOR INVESTIGATING THE IDEATIONAL CONTENT OF	235	Success Through Vocational Guidance (McKinney, J. and S.)	84
NEGRO CHILDREN, THE RELATION OF INTELLIGENCE TO AGE IN	378	SURVEYS, MENTAL, METHODS AND RESULTS OF	1
Notes Concerning Foreign Psychologists, The Psychological Corporation, Information Course on Science of Management and Applied Psychology	81	Systems, Psychotechnik und Taylor (Tramm)	217
Announcement of Change of Ownership of Journal of Applied Psychology, American Psychological Association, Second International Conference of Psycho-technics Applied to Vocational Guidance and to Scientific Management,	213	TAPPING TEST, STANDARDIZATION OF	113
PEARSON'S r FROM RANKED DATA, COMPUTATION OF	385	TEMPERAMENT, SPEECH AND WILL	291
PERSONAL TRAITS, RELIABILITY OF JUDGMENT OF	161	TEST SCORES, CONVERSION OF, INTO SERIES WHICH SHALL HAVE ANY ASSIGNED MEAN AND DEGREE OF DISPERSION, 298; TAPPING, STANDARDIZATION OF,	113
PERSONAL RESEARCH, SHIFT OF EMPHASIS NEEDED IN 141; Research, Agencies (U. S. Dept. of Labor, No. 299), 86; RELATIONS IN INDUSTRY	307	TESTIMONIAL AS AN ADVERTISING APPEAL	192
		TRUE FALSE EXAMINATION, THE INDIVIDUAL INJUSTICE AND GUESSING IN THE	342
		UNWILLINGNESS TO BE TESTED, SIGNIFICANCE OF	217
		VOCATIONAL GUIDANCE, Success Through (McKinney, J. and S.) 84; INTERESTS, METHOD FOR STUDY OF, 243; SELECTION, LETTER OF APPLICATION IN	74
		WILL-TEMPERAMENT, SPEECH AND	291
		Women Professional Workers (Adams)	83

INDEX OF AUTHORS

(The names of authors of original contributions are printed
in SMALL CAPITALS.)

ANDERSON, ROSE G.	1	KNIGHT, F. B.	211
ARLITT, ADA HART	378	KORNHAUSER, A. W.	302
BAKER, HARRY J.	349	LEWY-GUINZBURG, R. L.	221
BREED, F. S.	255	Link, Henry C.	219, 306
BROTEMARKLE, R. A.	235	LOWELL, FRANCES	276
BURKE, AGNES	205	Miner, J. B.	215
CHAPMAN, J. CROSBY	342	NACCARATI, SANTE	221
CLARK, RUTH SWAN	185	POFFENBERGER, A. T.	74, 156
COWDERY, KARL M.	311	REAM, M. J.	69
DOWNNEY, JUNE E.	291	REYMEET, MRS. ANNE	
FREYD, MAX	243	ROSS	34, 120
GARRISON, CHARLOTTE G.	205	ROOT, W. T.	267
GARTH, THOMAS R.	331	Sanford, Edmund C.	216
GOLDSMITH, DOROTHY B.	149	SLAWSON, JOHN	161
HEWES, AMY	113	THORNDIKE, EDWARD L.	29
HOLLINGWORTH, LETA S.	205	TOOPS, HERBERT A.	172
HULL, CLARK L.	298, 385	TOWN, CLARA HARRISON	89
JACKSON, A. W.	302	TURNER, EGBERT M.	192
JOHNSON, O. J.	403	VARTANIAN, V. H.	74
JONES, EDWARD S.	198	VITELES, MORRIS S.	391
KATZ, S. E.	255	WAGONER, LOUISA C.	291
KITSON, HARRY DEXTER	59,		
83, 84, 141, 216, 303, 307, 418			



The Journal of
Applied Psychology

Vol. VI

MARCH, 1922

No. 1

EDITORS

JAMES P. PORTER

Clark University

WILLIAM F. BOOK

Indiana University

And a Board of Co-operating Editors



CLARK UNIVERSITY
WORCESTER MASS.

FLORENCE CHANDLER, Publisher

Entered as second-class matter, June 9, 1917, at the post office at Worcester, Mass., under
No. 107, March 2, 1918.

Authorized by postoffice at special rate of postage provided for in section 1103, Act of
October 3, 1917. Authorized April 23, 1921.

Carnegie Institute of Technology; L. M. TERMAN, Leland Stanford University; E. L. THORNDIKE, Columbia University; I. E. WALLIN, Psycho-Educational Clinic, St. Louis; G. M. WHIPP, University of Michigan; HELEN THOMPSON WOODLEY, Vocational Bureau, Cincinnati.

The scope of the *JOURNAL OF APPLIED PSYCHOLOGY* will include following: (a) The application of psychology to vocational and such as law, art, public speaking, industrial and commercial work problems of business appeal. (b) Studies of individual mentalities as types of character, special talents, genius, and individual differences including the problems of mental diagnosis and vocational problems. (c) The influence of general environmental conditions, such as of weather, humidity, temperature; also such conditions as nutrition, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

VOLUME VI

MARCH, 1923

NUMBER

TABLE OF CONTENTS

ROSE G. ANDERSON.	Methods and Results of Mental Surveys.	1
EDWARD L. THORNDIKE.	On Finding Equivalent Scores in Tests of Intelligence.	11
ANNA ROSE REYNERT.	Some Factors of Aesthetic Judgment.	25
HARRY DEXTER KITSON.	Minor Studies in the Psychology of Advertising from Psychological Laboratory of Indiana University.	35
	V. The Development of Art-Forms in Magazine Advertising.	35
	VI. Color in Magazine Advertising.	45
	VII. Negative Suggestion in Advertising.	55
M. J. RYAN.	A Social Relations Test.	65
A. T. PROFFENBERGER AND V. H. VARTANIAN.	The Letter of Application in Vocational Selection.	75
NEWS AND COMMENT.		

Ε2
The Journal of
Applied Psychology

Vol. VI

DECEMBER, 1922

No. 4

EDITORS

JAMES P. PORTER

Ohio University

WILLIAM F. BOOK

Indiana University

And a Board of Co-operating Editors



PUBLISHED BY THE EDITORS

INDIANA UNIVERSITY PRESS

Bloomington, Indiana

Entered as second class mailing November 21, 1922 at the post office at Bloomington, Indiana, under the act of March 3, 1879.

Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917, authorized December 21, 1922.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. V. BINGHAM, Carnegie Institute of Technology; JUNE E. DOWNKY, University of Wyoming; ROBERT H. GAULT, Northwestern University; WILLIAM HEALY, Baker Foundation, Juvenile Court, Boston; H. L. HOLLINGWORTH, Columbia University; JOSEPH JASTROW, University of Wisconsin; HERBERT S. LANGFELD, Harvard University; JAMES H. LEURA, Bryn Mawr College; HENRY C. MCCOMAS, Princeton University; WALTER D. SCOTT, Northwestern University; C. E. SEASHORE, University of Iowa; DANIEL STARCH, University of Wisconsin; EDWARD K. STRONG, JR., Carnegie Institute of Technology; L. M. TERMAN, Leland Stanford Jr. University; E. L. THORNDIKE, Columbia University; J. E. WALLACE WALLEN, Teachers College, Miami University; G. M. WHIFFLE, University of Michigan; HELEN THOMPSON WOOLLEY, Vocation Bureau, Cincinnati.

The scope of the JOURNAL OF APPLIED PSYCHOLOGY will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeal. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

VOLUME VI

December, 1922

NUMBER 4

TABLE OF CONTENTS

CARL M. COWDERY, Measures of General Intelligence as Indices of Success in Trade Learning	311
THOMAS R. GARTH, Mental Fatigue of Mixed and Full Blood Indians	331
J. CROSBY CHAPMAN, The Individual Injustice and Guessing in the True False Examination	342
HARRY J. BAKER, Mental Tests as an Aid in the Analysis of Mental Constitution	349
ADA HART ARLETT, The Relation of Intelligence to Age in Negro Children	378
CLARK L. HULL, The Computation of Pearson's r From Ranked Data	385
MORRIS S. VITELES, A Comparison of Three Tests of General Intelligence	391
O. J. JOHNSON, Group Intelligence Examination for Primary Pupils	403
NEWS AND COMMENT	417
BOOK REVIEWS	429

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this JOURNAL; in every case the year, volume and pages should be specified.

THE JOURNAL OF APPLIED PSYCHOLOGY is published quarterly. The subscription price is \$4.00 a year. Single numbers \$1.25. Remittances and business communications should be addressed to the Editors, James P. Porter, Ohio University, Athens, Ohio or to William F. Book, Indiana University, Bloomington, Indiana.



Directory of American Psychological Periodicals

- American Journal of Psychology**—Ithaca, N. Y.: Morrill Hall.
Subscription \$6.50. 600 pages annually. Edited by Edward Bradford Titchener. Quarterly. General and experimental psychology. Founded 1887.
- Pedagogical Seminary**—Worcester, Mass.: Mary McLoughlin.
Subscription \$5. 575 pages annually. Edited by G. Stanley Hall. Quarterly. Pedagogy and educational psychology. Founded 1891.
- Psychological Review**—Princeton, N. J.: Psychological Review Company.
Subscription \$4.25. 480 pages annually. Edited by Howard C. Warren. Bi-monthly. General. Founded 1894.
- Psychological Bulletin**—Princeton, N. J.: Psychological Review Company.
Subscription \$5. 720 pages annually. Edited by Shepherd I. Franz. Monthly. Psychological literature. Founded 1904.
- Psychological Monographs**—Princeton, N. J.: Psychological Review Company. Subscription \$5.50 per vol. 500 pp. Ed. by James R. Angell. Published without fixed dates, each issue one or more researches. Founded 1895.
- Psychological Index**—Princeton, N. J.: Psychological Review Company. Subscription \$1.50. 200 pp. Edited by Madison Bentley. An annual bibliography of psychological literature. Founded 1895.
- Journal of Philosophy, Psychology, and Scientific Methods**—New York Sub-station 84. Subscription \$4. 728 pages per volume. Edited by F. J. E. Woodbridge and Wendell T. Bush. Bi-weekly. Founded 1904.
- Archives of Psychology**—Sub-station 84, N. Y.: Archives of Psychology. Subscription \$5. 600 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.
- Journal of Abnormal Psychology**—Boston: Richard G. Badger. Subscription \$5. 480 pages annually. Edited by Morton Prince. Bi-monthly. Entire field of abnormal psychology. Founded 1906.
- Psychological Clinic**—Philadelphia: Psychological Clinic Press. Subscription \$2.50. 288 pages ann. Ed. by Lightner Witmer. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.
- Training School Bulletin**—Vineland, N. J.: The Training School. Subscription \$1. 160 pp. ann. Ed. by E. R. Johnstone. Monthly (10 numbers). Psychology and training of defectives. Founded 1904.
- Journal of Educational Psychology**—Baltimore: Warwick & York. Subscription \$4. 540 pages annually. Monthly (9 numbers). Managing Editor, Harold O. Rugg. (Educational Psychology Monographs. Published separately at varying prices. Same publishers). Founded 1910.
- The Behavior Monographs**—Cambridge, Mass.: Emerson Hall. Subscription \$5. 450 pages per volume. Edited by John B. Watson. Published without fixed dates, each number a single research.
- Psychoanalytic Review**—Washington, D. C.: 3617 10th St., N. W. Subscription \$6. 500 pages annually. Ed. by W. A. White and S. E. Jelliffe. Psychoanalysis. Quarterly. Founded 1913.
- Journal of Experimental Psychology**—Princeton, N. J.: Psychological Review Company. Subscription \$4.25. 480 pages annually. Ed. by John B. Watson. Bi-monthly. Experimental. Founded 1916.
- Journal of Applied Psychology**—Indiana University Press, Bloomington. Subscription \$4. 400 pages annually. Edited by James P. Porter and William F. Book. Quarterly. Founded 1917.
- The Journal of Comparative Psychology**—Baltimore: Williams & Wilkins Co. Subscription \$5. 500 pages (approximately) per volume. Edited by Knight Dunlap and Robert M. Yerkes. Bi-monthly. Founded 1917.
- Journal of International Relations**—Worcester, Mass.: Louis N. Wilson. Subscription \$3. 500 pages annually. Edited by George H. Blakeslee and Harry E. Barnes, assisted by a Board of some 25 contributing editors. Quarterly. Founded 1910.



1

Stanford University Libraries



3 6105 006 683 481

STANFORD UNIVERSITY LIBRARIES
CECIL H. GREEN LIBRARY
STANFORD, CALIFORNIA 94305-6004
(415) 723-1493

All books may be recalled after 7 days

DATE DUE

AUG - 6 1997
AUG 01 1997

